

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 13:15:57 ; Search time 2611.63 Seconds  
(without alignments)  
16923.088 Million cell updates/sec

Title: US-09-826-206-1  
Perfect score: 2112  
Sequence: 1 ggcattggcgcgaatcatc.....tgaagaaatagggacgtg 2112

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
------------	-------------	-------	--------	-------	-------------

1	2112	100.0	2112	6	AX281573	AX281573 Sequence
2	2112	100.0	5859	6	AX281575	AX281575 Sequence
3	2007	95.0	2517	6	AX281576	AX281576 Sequence
4	2007	95.0	8509	6	AX281574	AX281574 Sequence
5	2007	95.0	8509	6	AX281868	AX281868 Sequence
6	49.2	2.3	81	6	AX019311	AX019311 Sequence
7	47.4	2.2	125020	9	AF429315	AF429315 Sequence
8	46.2	2.2	1969	6	AX024376	AX024376 Sequence
9	45.8	2.2	1969	6	AX024269	AX024269 Sequence
10	45.8	2.2	7218	6	166484	Sequence 14
11	44.6	2.1	32784	3	CE0865	770682 Caenorhabd
12	44.2	2.1	125020	9	AF429315	AF429315 Homo sapi
13	43.2	2.0	3490	4	BT414556	AJ414556 Bos tauri
14	42.8	2.0	13291	1	AE005932	AE005932 Caulobact
15	42.4	2.0	1141	6	AX083744	AX083744 Sequence
16	42.4	2.0	2325	3	HER278115	AJ278115 Helicobac
17	41.6	2.0	109489	2	AL669833	AL669833 Mus muscu
18	41.1	1.9	16488	1	AE004992	AE004992 Halobacte
19	41.1	1.9	44364	3	CELX67D2	AE004992 Caenorhab
20	40.8	1.9	10223	1	AE005476	AE005476 Escherich
21	40.8	1.9	296827	1	AP002561	AP002561 Escherich
22	40.4	1.9	731	6	E03296	E03296 cDNA sequen
23	40.4	1.9	11584	1	AE000333	AE000333 Escherich
24	40.4	1.9	16969	1	D90875	D90875 E.coli geno
25	40.4	1.9	74183	2	AC105624	AC105624 Rattus no
26	40.2	1.9	40.2	1	AX019312	AX019312 Sequence
27	39.8	1.9	10886	6	AX347036	AX347036 Sequence
28	39.8	1.9	138020	3	AC024776	AC024776 Caenorhab
29	39.8	1.9	309026	2	AC006760	AC006760 Caenorhab
30	39.6	1.9	41322	3	CEC50B6	281050 Caenorhabd
31	39.2	1.9	224577	2	AC097745	AC097745 Rattus no
32	39.2	1.8	59762	8	AB023032	AB023032 Arabidops
33	39.2	1.8	81704	6	AC094807	AC094807 Arabidops
34	39.2	1.8	119472	2	AC094164	AC094164 Rattus no
35	38.8	1.8	7141	14	AB018418	AB018418 Red sea b
36	38.8	1.8	40535	3	U50312	U50312 Caenorhabd
37	38.6	1.8	3146	8	AF118119	AF118119 Gracillari
38	38.6	1.8	8939	3	AF047652	AF047652 Caenorhab
39	38.6	1.8	18302	3	CEY2H9A	AF021448 Caenorhab
40	38.6	1.8	45817	2	AC006843	AC006843 Caenorhab
41	38.4	1.8	1998	8	ACY306973	AJ306973 Aegilops
42	38.4	1.8	34397	3	AF000198	AF000198 Caenorhab
43	38.4	1.8	77483	2	AC027171	AC027171 Homo sapi
44	38.4	1.8	157511	9	AC074348	AC074348 Homo sapi
45	38.2	1.8	3966	1	AF048718	AF048718 Rlmerrell

## ALIGNMENTS

RESULT 1  
AX281573  
LOCUS AX281573 2112 bp DNA  
DEFINITION Sequence 1 from Patent WO0177347.  
ACCESSION AX281573  
VERSION AX281573.1 GI:16608825  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (sites)  
AUTHORS D'Elia,J.  
TITLE Ketogulonigenium shuttle vectors  
JOURNAL Patent: WO 0177347-A 1 18-OCT-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
source  
1..2112  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="replicon from pADM291"

BASE COUNT	530 a	509 c	605 g	468 t
------------	-------	-------	-------	-------



Query Match	100.0%;	Score 2112;	DB 6;	Length 2112;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2112; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	ggatgagtgctgaataatcatagaatttggtagggagcgtacagggctctbacaaggagt	60
Db	1	ggcAAATGGGTGAAATTCATAGAAATTTTGTGTAGAGTCCGATGCGGCTCTACAGGGGGTG	60
QY	61	ctgcgcgagaga tctctggtctccagtagagcgacaatgagaggtgltagttgcctccctg	120
Db	61	CTCGCGGAGATCTCTGTCTCAGGTAGGGGACAAATGGAAGGTGTTAGTTGGCCCTCG	120
QY	121	tatgctctctgcgttggcgacattggtatctcttcctcccgacatctgattccctagag	180
Db	121	TATCCCTCTGTGCGGGGGCATGGGATATCTCTCCCGACATATGATATTCCTGCTAGC	180
QY	181	gattacatagattctctgcctctgcgggctctgtagggctgttcgggctgttcgggctgtgc	240
Db	181	GATTACTGATAGTTTCTGCTCTGTGCGGCTTGTGCGGGCTGTGTCGGGCTTGTCTGTC	240
QY	241	gggctgtccctctgtctgtccgcgctctccacttltacacataaaaaatggcggaagc	300
Db	241	GGGCTGTCCCTCTGTCTGTCCCGCTCTCTCTACATTTCATCAATAAAAAATGGGGAAGC	300
QY	301	ccctctctgtctatagttcttataagttacatacgaanaatlaacataatlaacatagctt	360
Db	301	CCCTCTGTCTTATAGTTCTTATAGTTTCATACGAAATAATTACAAATAATATCAATAGCTT	360
QY	361	atctgcttaaaaggagtaattggccgcgaagaaggagtaatttggccgcgaagaaggagt	420
Db	361	ATTTCGCTTAAAGGGAGTATTGGGCGCAAAAGGAGTATTGGGCCCAAAAGGAGT	420
QY	421	aattggccgcgaagaaggagtaatttggcggaatatcggttgttcaatgtagggagatcc	480
Db	421	AATTGGGCGCAAAAGGAGTATTGGGCGCAATCGGTGTTCATGCGGAGGAATCC	480
QY	481	cccttaatcatcttccccatgggaaagaacacaagttggccgcgaagccgggcttcgcac	540
Db	481	CCCTTAATCATTTCTCCCATATGGGAAAGACAAACAAGTGGCCGCACCGGCGCTTCGCAC	540
QY	541	cgagacaaaactgtgtccctcgtccgaggttggcgagaggggtctatatactgcgaatccgcc	600
Db	541	CNACAAAAACTGTGCTCCCTCGGCGAGTGGCGAGGGGCTATATGCGCAATCCGCCCC	600
QY	601	cgccgtcgaagcgctcaagctcaatgaatlaatgaatagccacttgcggcgccgcatgagct	660
Db	601	CGCCTTCAGGCGCTTCAAAGCTCATAGCTATTATATATAGCCACTGCGGGCCGCCATNGCT	660
QY	661	gatgatgtgctgcatagaatgtcggtctggccgcgaattcgcgaatctgaacgcatgaaaac	720
Db	661	GATGATGTGCGCATGAATATGCGGTGGCGGACATTCGCGCAATCGACGCGCATGAANAAC	720
QY	721	catgaacgtagaagcctgtgaccccgctgtttgagaagaacttagcgccttcggtgttgacccat	780
Db	721	CATGACCGTAGAGACCTCGACCCCGCTGTGTGAGGAGCTTACCCCTCGCGGTGTGACCCAT	780
QY	781	gatgaaccctcgaagaatgatcgtgtgacagcttcggcggtcttgttgcagtagagcgcgaaatagac	840
Db	781	GATGACCCCTGCAAAAGATGATGTGACAGTGTGGCGGCTGTGTGATGAGCGCGCAATAGAC	840
QY	841	taaccgcagaagcgaagcggcggaactcctctagtgaacgtggagaccttcggaataacttcggt	900
Db	841	TACGCGCAGGAGGACACGGGGAACCTCTATGTGACGTGAGACTTCGCGAGTATTCGCT	900
QY	901	cgtagtcggcgaggtcgaaccacttggcgacatcttcgacgcgtcaaaacggtatltccatctc	960
Db	901	CGTAGTCGGCGGAGTCAACACTTGGGCCATTTCTCAGCCGTCAAAAGGATATTCATCATCTC	960
QY	961	ggtagtaagtaattccgtgtgtgttctccagcagctctctagctctcgccaatcttgatcgg	1020
Db	961	GGTAGTAGTATTCGAGTCTGCTCTTCAGACGCTCTCTATCTTCGGCAACTTGTGATCGG	1020

QY	1021	atgagcgcgaaaacctttaaagtcctcccgagtttcggygcctctttagatgcgagga	1080
Db	1021	atgacgcggaanaacctttacgggtccccaagtgcggcgctctttagatgcgagga	1080
QY	1081	aagaatggttcgttgaacgagcgttaacagatttgcttcaacctgcactgcattgagatc	1140
Db	1081	aagatggttcgttgaacgagcgtttaacagatttgcttcaacctgcactgcattgagatc	1140
QY	1141	aacctttaaactgcgtctgaacatttcgcgcgaacgccaagaatttgcgcgttaagctgca	1200
Db	1141	aacctttaaactgcgtctgaacatttcgcgcgaacgccaagaatttgcgcgttaagctgca	1200
QY	1201	actgttactataatgcgtggaagtgaagaacgaacccaaacgcttcgcgaagcgagagctg	1260
Db	1201	actgttactataatgcgtggaagtgaagaacgaacccaaacgcttcgcgaagcgagagctg	1260
QY	1261	ggttccaagatgcgttcgagaatgcgttcgcgaagggcagcgcgaacagatagaccccttc	1320
Db	1261	ggttccaagatgcgttcgagaatgcgttcgcgaagggcagcgcgaacagatagaccccttc	1320
QY	1321	ttccccaagaagcgggcgggaatcaactaagttccaagtttgcttgcgtgagctgaagcttgc	1380
Db	1321	ttccccaagaagcgggcgggaatcaactaagttccaagtttgcttgcgtgagctgaagcttgc	1380
QY	1381	ggagagaacaaagacaacagatcgtacgtcctaagattccgcgcgttctgcgtcggagaga	1440
Db	1381	ggagagaacaaagacaacagatcgtacgtcctaagattccgcgcgttctgcgtcggagaga	1440
QY	1441	ggcgtgcgtctgagcgtctgaacatcgaacaaaaacgttcttgatcttcgcgcgaagga	1500
Db	1441	ggcgtgcgtctgagcgtctgaacatcgaacaaaaacgttcttgatcttcgcgcgaagga	1500
QY	1501	gggaaggttgagtttgaagatcttcaacgcgaatgctgttaatgaatttcgttgaagc	1560
Db	1501	gggaaggttgagtttgaagatcttcaacgcgaatgctgtgttaatgaatttcgttgaagc	1560
QY	1561	atgtgcataatagcgttgaagactatgaatatcacgcgtgcgaagcttcaaaaacgag	1620
Db	1561	atgtgcataatagcgttgaagactatgaatatcacgcgtgcgaagcttcaaaaacgag	1620
QY	1621	gttgtggcaccgcgaacccaatcaccgcggcgcttaaaagcgttaaaattccgcgttaaaaag	1680
Db	1621	gttgtggcaccgcgaacccaatcaccgcggcgcttaaaagcgttaaaattccgcgttaaaaag	1680
QY	1681	atgaatctggcgacatggttatagatccttcgcgaatgtcacagagtgttccctccatt	1740
Db	1681	atgaatctggcgacatggttatagatccttcgcgaatgtcacagagtgttccctccatt	1740
QY	1741	caaaagaatatcacccgaacacctcaacgcgaagatratgtttaagcgtgatgaaacacatg	1800
Db	1741	caaaagaatatcacccgaacacctcaacgcgaagatratgtttaagcgtgatgaaacacatg	1800
QY	1801	aaatgaacctcagaatacagcgcattlaagcgttgtaagtttcggaacttaacgcgatgcttat	1860
Db	1801	aaatgaacctcagaatacagcgcattlaagcgttgtaagtttcggaacttaacgcgatgcttat	1860
QY	1861	ctgatgcagaggagatctgcgaacaaatggcgcgacatggcggaagcgtcttcaattcat	1920
Db	1861	ctgatgcagaggagatctgcgaacaaatggcgcgacatggcggaagcgtcttcaattcat	1920
QY	1921	caccgatagaagagaagaacgcgcctccctcctcaaaaacaaagatgtgtgaagaatatctgat	1980
Db	1921	caccgatagaagagaagaacgcgcctccctcctcaaaaacaaagatgtgtgaagaatatctgat	1980
QY	1981	cctggagcttcagagagccttgccttactgcgcgaaaaaacgcgatattgagagcgaagccc	2040
Db	1981	cctggagcttcagagagccttgccttactgcgcgaaaaaacgcgatattgagagcgaagccc	2040
QY	2041	gaactttgagagcggaagcctataacagatgccaaaacactagaagccagatltgaagaaa	2100
Db	2041	gaactttgagagcggaagcctataacagatgccaaaacactagaagccagatltgaagaaa	2100
QY	2101	atcagggaacgtg 2112	



Db 2101 ATAGGAAACGTG 2112  
RESULT 2  
AX281575 5859 bp DNA linear PAT 02-NOV-2001  
LOCUS AX281575  
DEFINITION Sequence 3 from Patent WO0177347.  
ACCESSION AX281575  
VERSION AX281575.1 GI:15660827  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial sequence.  
REFERENCE 1 (sites)  
AUTHORS D'Elia, J.  
TITLES Ketogulonigenium shuttle vectors  
JOURNAL Patent: WO 0177347-A 3 18-Oct-2001;  
Archier-Daniels-Midland Company (US); D'Elia, John (US)  
FEATURES  
source 1. 5859  
location/Qualifiers  
BASE COUNT 1444 a 1487 c 1566 g 1362 t  
ORIGIN  
Query Match 100.0%; Score 2112; DB 6; Length 5859;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ggcgaatggtcgaaatcagaatttctgtgaggtgctgagcgcttcagacggggtg 60  
DB 402 GGCATAGGCTCGAATTCATGATTTTGTGTAGGTGCTGAGGCTCTGACAGGGGTG 461  
QY 61 ctgcgcgagatctctgtgtctcaggtagggcgacaatgagaggtgttgaattgcccctg 120  
DB 462 CTGGCGGAGATCTCTGCTGCTCAGGTAGGGGACAAATGAGAGGTGTAGTGGCCCTG 521  
QY 121 tctgcctctctgtgtgagcaattgggtcatctgctgcccagacatatgatctcgctag 180  
DB 522 TATCGCTCTGCGTGGCGGATGGGTCAATCTGCCGACATATGATATTCCTCTAGAG 581  
QY 181 gattactagatagtttcgctcgttcgtgagcttgcgagcttgcgagcttgcgagctgtc 240  
DB 582 GATTACTAGATAGTTTCTGCTGTGCGCTTGTGCGGCTTGTGCGGCTTGTGCGGCTTGTG 641  
QY 241 gggcctgtccctctgtccgcgcgtctcacttttcaacaataaataatgagcgagc 300  
DB 642 GGGCTGTCCCTCTGTGTCGGCGCTGTCTCACTTTTCACAAATCAAAAATGGCGAAGC 701  
QY 301 cctcttctgtctatagttctctatagttcacaagaataatacaataatcaatagctt 360  
DB 702 CCTTCTGTCTATAGTTCTTATGTTCAACGAAATTAACATATTAATCAATAGCTT 761  
QY 361 attcgcctaaaggagtaatttggccgcaaaaggagtaatttggccgcaaaaggagtg 420  
DB 762 ATTGCTTAAGGAGATTAATGGCGCAAAAGGAGTAATTTGGCGCAAAAGGAGGT 821  
QY 421 aattggcgcaaaaggagtaatttggccgcatatcgttcttaccatgtyggagaaatcc 480  
DB 822 AATTGGGCGCAAAAGGAGTAATTTGGCGCATATCGGTTTACATGGGAGGAGTAATCC 881  
QY 481 ccttaatacatttctcccatgggaagaacaacaagaattggccgacagacgggcttcgac 540  
DB 882 CATTAAATCAATTTCTCCCATGGGAACAACAACAAGTGGCGCAGACCGGCTTTCGAC 941  
QY 541 cagacaaatctgtctccctgaggaattggcgaggtgaggtctatattgcaaatccgccc 600  
DB 942 CAGACAAATCTGTCTCTCCGCGGAGGTGGCGAGAGGGGTCTATATGCGCAATTCGGCC 1001  
QY 601 cgctgtcagcgctcaagatcagatcatttaataagtaagcaatcgcgcgagctatggct 660

Db 1002 CGCCTGACGGCGCTCAAGCTCATGCATTTAATGATAGCCACTGCGGGCGCGCATGGCT 1061  
QY 661 gatgatgtgcgccatgaaatgctggtcgccagacttcgccaatccagagcatgaaaaac 720  
DB 1062 GATGATGTGCGCCATGAATGCGGTGCGCCAGATTCGCGCAATTCAGAGGATGAAAAAC 1121  
QY 721 catgacgctgagagccctgagaccgcgtgtctgaggaagctagcgcgctgtgttaccat 780  
DB 1122 CATGACCGTGAGACCTGACCCCGCTGTTCAGAGAGCTAGCCGCTGCGGTGTGACCAT 1181  
QY 781 gatgaccctgcaaatgatgctgacagctcggcggtctgtgtcgtatgaagcgcaaatagac 840  
DB 1182 GATGACCCCTGCAAGATGATGATGACATGCGGCGCTGTGTGATGAGCGCGAATATAGC 1241  
QY 841 taaccgcaagagagcaagcgcggaactcctagtgaagctgtgacccctccgagatcactcgt 900  
DB 1242 TACGCGCAAGAGGAGCAAGCGGCAACTCTAGTGAAGTGAAGCTTCGAGATACATTCCT 1301  
QY 901 cgtatgcgcgagagctcgaaccactgaggcattctcgaacgtcaaacgtattccatctc 960  
DB 1302 CGTATGGCGCGGAGTGCACACCTAGGCGCATTTCTGACCCGTCAAAACGGTATTCCATTC 1361  
QY 961 gtagtaagtaattccgtgtgtctgttccagacgctctcctagctcgcgaatctgtatcgt 1020  
DB 1362 GGTAAGTAAGTATCCGTCGTCCTGTTCACAGCAGCTAGTCTGCGCAATCTTGATCGG 1421  
QY 1021 atgagcgcgaaaccccttaagtgctcccgagcttgcggcgctcttgagtgcccgagga 1080  
DB 1422 ATGAGCGCGAAACCTTTACGCTGCCCGAGTGGCGGCGCTTGGAGTGTCCCGAGGA 1481  
QY 1081 aagaatgtctgtgaaacgagcttaacagatttctctcaaacctgcagctgatagatc 1140  
DB 1482 AAGATGCTTCTTGAACGACGTTTACGATTTGCTCTCAAACTGCACTGATGATGATC 1541  
QY 1141 aaccattatcgcgtctcagacatgacgycgcaaacgcaaatgctgagctgtgagca 1200  
DB 1542 AACATTATATCGCGTCAATTCGACGCAAAAGCGCAAGATTTGGCGTAGCGTGCA 1601  
QY 1201 agtgtgactaagcttgggaagtgaagaagacccaacgctcgcgcagcgcgagctgtgcg 1260  
DB 1602 AGTGTGACTATAGCTGTGGAAGTGAAGACACCCCAACCGTCGCGAGCGGAGTGTGCG 1661  
QY 1261 ggttccaaagtcggtctgagagctgtcgtcgaagagggagggaggaagataagcccccc 1320  
DB 1662 GGTTCCAAAGTGTGCTGAGATGCTGTGCGAGAGGGGAGGAGGAGAACATAGCCCCCTCC 1721  
QY 1321 ttcccaagaagcgcgagatcaccatcagctacagcttgcgtgagagctgtaaacgctctgtc 1380  
DB 1722 TTCCAGAGCGGGGGGATCATCACTACAGTCCAGCTGTGGTGGAGACTGAACGCTCTGCT 1781  
QY 1381 ggcagcaacaagaacagatctgagctcgaagctcgcgagcttctctgcggagagaga 1440  
DB 1782 GGCAGCAACAAGGACCAAGATCTGTGCCCTCAGACTTCGCGCGCTTGTGCGGAGAGA 1841  
QY 1441 ggcgtgcgtctgagcgctgcgaacaatcgaaaaacgcttttagatcttcgtgcgaaggt 1500  
DB 1842 GCGGTGCTGTGACGCTGCAGCAACATCGAANAACGTTTATGATTTTGTGCGCAAGGTA 1901  
QY 1501 gggaaggtttagattttaggtlatcttcacgcaatagtgcttaaatgacttctgtgaaacg 1560  
DB 1902 GGAAGAGGTTTGAAGTTTGAAGTATTTACCGCAATAGTGAATGACTTTTCGTGAACG 1961  
QY 1561 atgtgcaataaagcgtgtaagactatgaatacaagcgctgagacagctgcgaagaagcag 1620  
DB 1962 ATGTGCAATATAGCGGTAAAGCTATGAATACACGCGTGGAGAGGCTCCAAAAGCAAG 2021  
QY 1621 gtgtgagcaacgcaaccaatcactcgtggtgcctaaagaaggttaaaaatttcggtlaaaaag 1680  
DB 2022 GTGTGCGACCGCAACCTACCTGCGGCGCTTAATAAGCGGTAAATTTCCGGTAAAAAG 2081  
QY 1681 atgaatctggggcattgggtlatagatctcagaaattgcaagaagtgcttctcccatct 1740



```

Db 2082 ATGAACTGGGGCATGGTTATAGATCTGCAGAAATGCGACAGAGTGTTCCTCCCATTT 2141
Qy 1741 caaagaataacacgaacactaacacgaatataatgtaagctgtaagaacacatg 1800
Db 2142 CAAAGAAATACACCGAAACCTTAACACGCAAGTATATGTAGGCTGATGAACACATG 2201
Qy 1801 aatagccctcagaataacagcgcattagagcgtgaagltcgcgacttaacgcgactta 1860
Db 2202 AATGACCTCAGAAATCAGCGCATTTAGACGCTGAAGTTGCGACTTACCGCATTTAT 2261
Qy 1861 ctgagtcacaggaagatcgcgcacaaatgacgcacatgacgcgcgcgcgcgcgcgcgc 1920
Db 2262 CTGATGCCAGGAGGATCCGACAAATGCGCACATGCGCGACGCTCTTCAATTTCAT 2321
Qy 1921 caccgctgagagaggaagacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1980
Db 2322 CACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2381
Qy 1981 cctggagctcagagagccttgcttactgagcgcgcgcgcgcgcgcgcgcgcgcgcgc 2040
Db 2382 CCGGGGCTTACGAGGCTTCTTACTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441
Qy 2041 gcaacttaagagagagagagcctataacagatraccaaaacactagaagcagagaa 2100
Db 2442 GCACTTTAGAGGCGGAGAGCTTATACGATACCAAAACACTAGAGCCAGATTGAGGAA 2501
Qy 2101 atagggaacgtg 2112
Db 2502 ATAGGAAAGCTG 2513

RESULT 3
AX281576 2517 bp DNA linear PAT 02-NOV-2001
LOCUS AX281576 Sequence 4 from Patent WO0177347.
DEFINITION AX281576
ACCESSION AX281576
VERSION AX281576.1 GI:1660828
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (sites)
AUTHORS D'Elia, J.
TITLE Ketogulonigenium shuttle vectors
JOURNAL Patent: WO 0177347-A 4 18-OCT-2001;
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)
FEATURES
source 1. .2517
/organism="synthetic construct"
/db_xref="taxon:32630"
/Note="Ketogulonigenium part of pADM291-4"

BASE COUNT 657 a 589 c 683 g 588 t
ORIGIN
Query Match 95.0%; Score 2007; DB 6; Length 2517;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ggcacatggatcgaaatcattagaattctgtgaggtgctgtagcgcctcgcagaggtg 60
Db 1 GGCAAATGGCTCGAAATTCATAGAAATTTGTGTAGGCTGTGAGCGGCTGTGACAGGGGtg 60
Qy 61 ctggcgagagatctctgtgctcaggtagggcgacaatgtagaggtgttagttgccccctg 120
Db 61 ctggcgagagatctctgtgctcaggtagggcgacaatgtagaggtgttagttgccccctg 120
Qy 61 ctggcgagagatctctgtgctcaggtagggcgacaatgtagaggtgttagttgccccctg 120
Db 61 ctggcgagagatctctgtgctcaggtagggcgacaatgtagaggtgttagttgccccctg 120
Qy 121 taccgtctctcgtgctgagcattggatcctgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 121 taccgtctctcgtgctgagcattggatcctgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Qy 121 taccgtctctcgtgctgagcattggatcctgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 121 taccgtctctcgtgctgagcattggatcctgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Qy 181 gattactgaagttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 240
Db 181 gattactgaagttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 240

```

```

Db 181 GATTACTGATAGTTTCTGCTCCCTGTCGGGCTTCGCGGCTTGTGCGGCTTGTGCGGCTTGTG 240
Qy 241 gggcctgtccctctgtccgcgcgtccctcaacttttacacataaaatggcgagac 300
Db 241 GGGCCTGTCCCTCTGTGCTCCCTGTCGCTTCTCTCACTTTTACAAATGAAATGGCGAAGC 300
Qy 301 cctctgtctctatagttcttcttaagttcacgaaataacataatcaatcaatgact 360
Db 301 cctctgtctctatagttcttcttacttacttacttacttacttacttacttacttacttact 360
Qy 361 attcgtcttaaaaggagtaattggcgcaaaaaggagtaattggcgcaaaaaggagtaattggcg 420
Db 361 ATTGCGCTTAAAGGAGATATTGGCGCGCAAAAGGAGTAATTGGCGCGCAAAAGGAGT 420
Qy 421 aattggcgcaaaaaggagtaattggcgcaaaaaggagtaattggcgcaaaaaggagtaattggcg 480
Db 421 AATTGGCGCGCAAAAGGAGTAATTGGCGCGCAAAAGGAGTAATTGGCGCGCAAAAGGAGT 480
Qy 481 ccttaatactctccctccttggaagacaacaagaatgagcgcgcgcgcgcgcgcgcgcgcgc 540
Db 481 CCTTATCATTTTCTCCCAATGGGAAAGACACAAAGTGGCGCGAGACCGGCTTCGAC 540
Qy 541 caagacaaaactgtgctccctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db 541 CAGACAAAACCTGTCTCTCCCTGCGAGGTGCGAGAGGGGCTTATATGCGCAATCGGCC 600
Qy 601 cgcctcagagcgtcctcaagctcagctcagctcagctcagctcagctcagctcagctcagct 660
Db 601 CGCCTGACGGCGCTCAAGCTCATGCAATTAATGATAGCCATGCGGCGCGCGCATGGCT 660
Qy 661 gatgatgtgcgcacatgaatgtgcgtgctgctgctgctgctgctgctgctgctgctgctgct 720
Db 661 GATGATGTGCGCCATGAATGTGCGGTGCGGCGCATGTCGCGCAATGCGCAATGCAAAAC 720
Qy 721 catgacgctgagagcgtcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
Db 721 CATGACCGCTAGAGACCTGACCCCGCTGTGAGAGAGTACCCCTGCGGCTGTGACCAT 780
Qy 781 gatgacccctgcaaatgatgatgatgatgatgatgatgatgatgatgatgatgatgatgatgat 840
Db 781 GATGACCCCTGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 841 taccgcaagagagagagagagagagagagagagagagagagagagagagagagagagagag 900
Db 841 TACCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 901 cgtatggcgagagagagagagagagagagagagagagagagagagagagagagagagagag 960
Db 901 CGTATGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 961 ggtagtaagatattcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1020
Db 961 GGTAGTAAGATATTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Qy 1021 atgacgcgcaaaacctttacggtccgaggtggcgaggtggcgaggtggcgaggtggcgaggt 1080
Db 1021 AATGATGCTTCTGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 aagatggtcgtgctgagcagcgttaacagattgcttcccaacctgacactgacactgacact 1140
Db 1081 AAGATGCTTCTGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Qy 1141 aaccattatcgcgcgtcgtacattgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1200
Db 1141 AACCATTTATCCGCTGTGACATTTAGAGCGCAAAAGCGAGACAGATTTGCGGTAGCGTGCA 1200
Qy 1201 agtgtactatagagcgtggaagtgaagaagcgaacacgcgcgcgcgcgcgcgcgcgcgcgcgc 1260
Db 1201 AGTGTACTATAGAGCTGGGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1261 ggttcaaggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1320
Db 1261 GGTTCAGAGGTGCTGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320

```



Qy 1321 ttcccaagaacgagcgagatcaaccatacagcttggctggagctgaaacgctctgct 1380  
|||||  
Db 1321 TTCCCAAGAAGCGGCGGAGATCACTACAGTTCAGCTTGGCTGGAGAGCTGAACGCTCTGCT 1380  
Qy 1381 ggcagcaacaagaacaaacgactgacgctcagactccgagcttctcgtcggagaga 1440  
|||||  
Db 1381 GGCAGCAACAAGAACAGATCTGATCGCCTCAGACTTCCGCGCTTCTGTCGGAGAGA 1440  
Qy 1441 ggcgtgctctgagcgtctgcaaaatcgaataaactgtttttaagatttctcgtcgaagta 1500  
|||||  
Db 1441 GGCCTGCTCTGAGCGCTGCAAAACATGTAAGTCTTTTAAATTCCTGGCAAAAGTA 1500  
Qy 1501 gggaaagcttgaagcttgaagcttcaaccgcaatagtgtaaaatgacttcgtaaacg 1560  
|||||  
Db 1501 GGGAAAGCTTGAAGCTTTGAGAGTATTCACCCGCAATAGTGAATGACTTTCGTGAACG 1560  
Qy 1561 atgtgcaataatagcgtgaagactaagaatacagcgtcgaagcgtcgaacaaacgag 1620  
|||||  
Db 1561 ATGTGCAATATAGCGGTAAAGACTATGAATAACGCGCTGACAGGCTGCAAAAGCAACGG 1620  
Qy 1621 gttgtggagcgcgaacacatcaactcggcgctaaacgagtaaaatttcggtataaaag 1680  
|||||  
Db 1621 GTGTGGGACCGCAACCATCACTCGGCGCTAAAGGCGTAATAATTCCGGTAAAAAAG 1680  
Qy 1681 atgaatctggagcgtggttataagactcgtgagattgcaagagtgcttcctccattt 1740  
|||||  
Db 1681 ATGAATCTGGGGCATGGGTTATAGATCTGTGAGAAATTCACAGAGTGTTCTCCCATTT 1740  
Qy 1741 caaagaatacacaccgaacacactaacacagcaaglatatgtaagcgtgatagaacacatg 1800  
|||||  
Db 1741 CAAAGAATATACCCGAAACACCTTAACAGCAAGATATATGTAAGCGTATGAACACATG 1800  
Qy 1801 aaatgaactcaagaatacagcgtcatagaagcgtgaagcttcgagacttcaagcgtcttat 1860  
|||||  
Db 1801 AAATGACCTCAGAAATAGCGCATTTAGAGCGTGAAGTTCCGACTTTACGCATCTTTAT 1860  
Qy 1861 ctgagtcgaaggaagatcgcagcaaatcgcgcgacatgagccgagcgtcttcaattcat 1920  
|||||  
Db 1861 CTGATGGCAGGAGAGATCGCGACAAATGCGCGACATGCGCGACGTTTCAATTTGAT 1920  
Qy 1921 caaccgatgagagaagaacaccccccctcaaaaaaagaatgltggaagatalctgcat 1980  
|||||  
Db 1921 CACCGATGAGAGAGAGAACCGCCCTCCCTCAAAAACAAAGATGTTGMAAGATATTCGAT 1980  
Qy 1981 cctggagctcagagagccttgccttactatgagaga 2015  
|||||  
Db 1981 CCTGGCTCAAGAGCCTTGCTTTAAAAACCTGAA 2015

RESULT 4  
AX281574 8509 bp DNA linear PAT 02-NOV-2001  
LOCUS AX281574  
DEFINITION Sequence 2 from Patent WO0177347.  
ACCESSION AX281574  
VERSION AX281574.1 GI:1660826  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct.  
artificial sequence.  
REFERENCE 1 (sites)  
AUTHORS D'Elia,J.  
TITLE Retroviral shuntle vectors  
JOURNAL Patent: WO 0177347-A 2 18-0CT-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
Source location/Qualifiers  
1..8509  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="pADM291"  
BASE COUNT 2065 a 2239 c 2345 g 1860 t  
ORIGIN

Query Match 95.0%: Score 2007; DB 6; Length 8509;  
Best Local Similarity 99.8%: Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 ggcagtggtcgaatcatagaaatttctgtgagtgctgtagcggctctgacagagtg 60  
|||||  
Db 2955 GGCAATGGCTGGAATATTCATGAAATTTTGTGTGAGGTGCGTATGAGCGCTCTGACAGGGTG 3014  
Qy 61 ctgcgcgagatctctgtctcgaagtaggcgacaaatgagaagtgcttgcctctg 120  
|||||  
Db 3015 CTGGCGGAGATCTCTGTCTCAGGTAGGGACAAATGAGAGGATGATGTTGCCCTG 3074  
Qy 121 tctgcctcgcgtgagcattgagatcctgcccagacatatgatatccgttagag 180  
|||||  
Db 3075 TATGCTCTTCTGCGGGCGCATTTGGGTATCTCGCCGACATATGATTTCCGTTAAG 3134  
Qy 181 gattactgatatgtctcgtcgtcggagctgtcggagctgtcggagctgtc 240  
|||||  
Db 3135 GATTACTGATAGTTTCTGCTGCTGCGGCTTGTCGGGCTTGTGCGGCTTGTGCGGCTTGTG 3194  
Qy 241 gggcctgtccctctgtctcgcgctgtctcacttctcaacatacaaaatggcggaac 300  
|||||  
Db 3195 GGGCGTGCCTCTGCTGCTGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCTTGTG 3254  
Qy 301 ccttcgtctcatagatctctatagttcatatagatacgaataatcatatcatatagctt 360  
|||||  
Db 3255 CCTTCTTCTTATATGTTCTTATATGTTATATGTTATATGTTATATGTTATATGTTAT 3314  
Qy 361 attcgtctaaagagagtaatttggcgcgaacaaaggaaglaatttggcgcgaacaaagga 420  
|||||  
Db 3315 ATTGCGCTTAAAGGAGATATTTGGGCCCAAAAGGAGTAAATTTGGGCCCAAAAGGAGT 3374  
Qy 421 aattgggcgcgaacaaaggaagtaatttggcgcgaatcgttctgttcaatggggaagatcc 480  
|||||  
Db 3375 AATTGGCGCGCAAAAGGAGATATTTGGGCCCAAAAGGAGTAAATTTGGGCCCAAAAGG 3434  
Qy 481 ccttaatcttctccccaatggaagaacaacaagaatggcgcgaacgcggagcgtctgac 540  
|||||  
Db 3435 CCTTATATATTTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3494  
Qy 541 cagaacaaacatgtgtctcctcgcgagtgagcgaagaggtatataatgacatccgcc 600  
|||||  
Db 3495 CAGACAAAAATGTTGCTCCTCGCCGAGGTGCGAGAGGGGCTATATGCGCAATTCGCCCC 3554  
Qy 601 cgcctgcagagcgtcgaatcatatgatttaataagcactcgtggcgagcgcgaatgct 660  
|||||  
Db 3555 CGCCTGCAGGGGCTCAACCTCATGATTAATGATGATGATGATGATGATGATGATGATG 3614  
Qy 661 gatgagtgcgcaatgaatgagcgtgagcgaatctgcgaatcgcgaatcgcgaatcgcga 720  
|||||  
Db 3615 GATGATGTCGCCCATGAAATGCGGCTGCGCCGACATTCGCCGCAATGAGCGATGAAAAAC 3674  
Qy 721 catgacgtgagagcgtcgaaccccgctgtcgaagagctagccgcgtcgtgttgaacat 780  
|||||  
Db 3675 CATGACGCTGAGAGCGCTGACCCCGCTGTTGAGAGAGTACGCGCGCTGTTGACCAT 3734  
Qy 781 gatgacctctgaaagatgatactgtagcaagctggcgcttgctgcatgagagcggaataac 840  
|||||  
Db 3735 GATGACCTTGCAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3794  
Qy 841 taccgccaagaggaacagcggaactcctagtagaagcttgcgaatcgcgaatcgcgaatcgc 900  
|||||  
Db 3795 TACCGCCAGAGAGGAGGCGGCAATCTTATGAGAGTGTGACTTCGGAGATCAATTCGCT 3854  
Qy 901 cgtatbgcggcgagtgcaacacactggcgaatctcgaacgltcaaacggtatccatctc 960  
|||||  
Db 3855 CGTATGCGGGGAGTCCAAACACATGCGCCATTTCTGACACCGCAAAACGGTATTCATCTC 3914  
Qy 961 ggtgtaagtatctcgtctcgtctgtctcagacagcgtctatgctcgaatcttgcgtg 1020  
|||||  
Db 3915 GGTAGTAAATATCCGCTGCTGCTTCCAGACAGCTCTATGATCTGCAACATCTGATCGG 3974



QY 1021 atgagcgcaaaacccctttagcgtcccgaggttggcgcgctcttgagtgcccgagga 1080  
|||||  
Db 3975 ATGAGCGCAAAACCTTTACGGTCCCGAGTTGCGGGCCCTCTTGAAGTCCGAGGGA 4034  
QY 1081 aagatgttcgttgaagacgtttaacagatltgtctcaaaccttcacatgtatgagtc 1140  
|||||  
Db 4035 AAGATGTTGCTTGGAGACAGCTTAACAGATTTCCTCAACCTCGACATGGATGAGATC 4094  
QY 1141 aacattatcgcttcgcatatgacgcgcaaacgcaagaatgtgcgttagcgttgaca 1200  
|||||  
Db 4095 AACCATTTATCGGCTGTGACATGAGCGCAAAACCCAGCAAAATTTGCGGTACCGTGGA 4154  
QY 1201 atgttacctatagctgtgggaagtgaaagaccccaacctgtccagcgcgagctgagc 1260  
|||||  
Db 4155 AGTGTACTTATAGCTTGGAGTGAAGAGACCCCAACCTCCGCGAGCGCGAGCTGGCG 4214  
QY 1261 ggttcaaaagtcgctgcagatgtcgtcgcaagggcgcaagcgaataagccccctcc 1320  
|||||  
Db 4215 GGTTCAGAGTGGTGGAGATGCTCGCAGAGGGGCAACGAAAGATAGCCCTCC 4274  
QY 1321 ttcccaagaagcggtgggtacactacatgccaagtgttggtgagtgagtgagcttgct 1380  
|||||  
Db 4275 TTCCCGAAGACGGGCGGAGTACCTACATGTCACGTTGGCTGAGCTGAAGACCTGCT 4334  
QY 1381 ggcagcaacaagaacacagatctgcatcgctcagacttccggttctgtcgagagaga 1440  
|||||  
Db 4335 GGCAGCAACAAGACACAGATCTGATGGCTCAGACTTCGGGGTTCTCTCGGAGAGA 4394  
QY 1441 ggcgtgcgttcgagcgtgcgaacatcgaaaaactgtttttagatltctgcgcaagata 1500  
|||||  
Db 4395 GCGGTGCTGTGAGCGCTGCMAACATCGAAMAACTGTTTAAATTCCTCGCAAAAGTA 4454  
QY 1501 gggaaagttggaatttttgaagtatttcaacgcaatagtgtaaatgtacttgcgaaagc 1560  
|||||  
Db 4455 GGGAGGTTTGAAGTTTGAAGTATTTCACCGCAATAGTGTAAATGACTTTCCGGAACG 4514  
QY 1561 atgtgcataatagcgttaaacatataacacggttgacagagctgcgaagaacagc 1620  
|||||  
Db 4515 ATGTGCATATAGGGGTAAACATATGAATTAACAGCGCTGGAGCGTGCAGAAACACACG 4574  
QY 1621 gttgtgagcgcgaacacatctcgtggcgcttaaaagcggttaaaatttcggttaaaaag 1680  
|||||  
Db 4575 GTGTGGGACCGCAACCATCACTCGGGCGCTAAAMAAACGGGTAAATTTCCGGTAAAG 4634  
QY 1681 atgaatttggggatgggttatagatctcagaaattgcagagatgtttccctccatt 1740  
|||||  
Db 4635 ATGAATCTGGGGCATGGGTTATAGATCTCTGAGAAATGCAAGAGTGTTCCTCCATT 4694  
QY 1741 caaagaatacacccgaacacacttaacacagcaagtataatgtaagcgtgatgaacacatg 1800  
|||||  
Db 4695 CAAGAATAATACCCGAAACACCTTAACACGCAAGTATATGTAAACCGTATGAAACACATG 4754  
QY 1801 aaatgacctagaataatagcgcattagaagcgtgaagtgcgaacttgcgagatcttat 1860  
|||||  
Db 4755 AAATGACCTCGAATAAGCGCATTTAGCGCTGAAGTTTCGACTTTACGCGATGCTTTAT 4814  
QY 1861 ctgtatgcagaggaatgcgcgaacaatggcgagcatggtcgagagcgttcttcaattcat 1920  
|||||  
Db 4815 CTGATGCGCAGGAGAGATGCGCAAAATGGCGCAGCATGGCGAGCGTTCCTCAATTTCAT 4874  
QY 1921 caacgatgagaggaagaacgcgcctcccaaaaaaagaagtgtgtgaagatatctgcat 1980  
|||||  
Db 4875 CACCGATGAGAGAGAAACCGCCCTCCCTCAAAAACAAAGATGTGGAGATATTCGAT 4934  
QY 1981 ccttgagctcagagagccttgcttactacgtcgagaa 2015  
|||||  
Db 4935 CCTGGGCTTCAGAGACCTTGCCTTTAAACCTGAA 4969

RESULT 5  
AX281868 8509 bp DNA linear PAT 02-NOV-2001  
LOCUS AX281868 8509 bp DNA linear PAT 02-NOV-2001  
DEFINITION Sequence 1 from Patent WO0177139.

ACCESSION AX281868  
VERSION AX281868.1 GI:16609119  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source Location/Qualifiers  
BASE COUNT 2065 a 2239 c 2345 g 1860 t  
ORIGIN

Query Match 95.0%; Score 2007; DB 6; Length 8509;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggcattggttcgaatcatagaaatttgtgtgagtgtagcgtctgacaggggtg 60  
|||||  
Db 2955 GGCATGCGTGCAGAAATTCATAGAAATTTGTGTGAGTGTGCTGACGCTCGACAGGGGTG 3014  
QY 61 ctgcgcgagaaatctctgtgtctcaaggttagggcgcaaatggagaggtgttagttccccctg 120  
|||||  
Db 3015 CTGCGCGAGATCTCTGGTCTCAGTGAAGGCGCAATGAGAGGTTTAATTGCCCCCTG 3074  
QY 121 tatcgctctgcgtgcgcatgggtgcacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180  
|||||  
Db 3075 TATCGCTCTGCGTGGCGCATTTGGTCACTCCCGGACATATGATATTCGCTAGAG 3134  
QY 181 gattactagatgttctgcgtctgcgtgtgcggctgtgcggctgtgcggctgtgcggctgtgc 240  
|||||  
Db 3135 GATTACTGATAGTATTGCTGCTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCTG 3194  
QY 241 gggcgtctccctctgtcccgccgctgtctacttcttcaacaatcaaaaagggcgagc 300  
|||||  
Db 3195 GGGCTGTCCCTCTGTTGCCGCTGTCTCACTTTTCAATCAATCAATCAATCAATCAATCAAT 3254  
QY 301 cctctctgtctatagttcttatagttcatcagaaaaatcaacataatltcaatagctt 360  
|||||  
Db 3255 CTTCTTGTCTATAGTCTTATAGTTTATAGTTCATACGAAATTAACATATTAATCATAGCTT 3314  
QY 361 attcgtctaaaggaggaatttggcgcaaaaggaggaatttggcgccgcaaaaggagat 420  
|||||  
Db 3315 ATTGCTTAAAGGAGTAAATTTGGCGCAAAAGGAGTAAATTTGGCGCAAAAGGAGT 3374  
QY 421 aattggcgccgaaggaggaatttggcgcatatctggttcttaccatgttggaagaaatcc 480  
|||||  
Db 3375 AATTGGCGCGCAAAAGGAGTAAATTTGGCGCATATTCGTTTTCATATGGGAGAAATTC 3434  
QY 481 ccttaatcatcttcccatatggaaagaacaacaaagtggcgcaagacggcgcttgagc 540  
|||||  
Db 3435 CCTTAATCATTTCTCCCATAGTGAAGACACAAAGTGGCGCAGACCGGGCTTGTGAC 3494  
QY 541 cagacaaaactgtgtctccctgcgaggttgcgaggaaggtgtctatagcgcaatccgccc 600  
|||||  
Db 3495 CAGACAAAACCTGTGCTCCGCGCAGAGTGGCGAGAGGGGTCTATATGCGCAATCCGCC 3554  
QY 601 cgcctgcagcgcccaagctcatgcatttaatgtaagcaactggtggcgcgagatggct 660  
|||||  
Db 3555 CGCCTGCAGCGGCTCAAGCTATGATTTATGATGACCATTTGGGGGGCGCGCATGGCT 3614  
QY 661 gatgatgtgcacatgaatgtgcgtgtgcgcacatctgcgcgaatcgacgcatgaataaac 720  
|||||  
Db 3615 GATGATGTGGCGCATGAATGAGGAGGCTGCGCATTCGCGCAATCGACGCGATGAAAAAC 3674



Qy 721 catgacggtgagagcctgagccccgcgtgtctcgaggagctagccgctgctgtctgacccat 780  
Db 3675 CATGACGGTGAAGACCTGACCCCGCTGTTGAGAGAGCTAGCGCGTGGCTGTGACCAT 3734  
Qy 781 gatgacctgcaagaatgatactgtgacagtcgaggcttgctcgatgagcgcgaaatagac 840  
Db 3735 GATGACCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3794  
Qy 841 taaggcgaag 900  
Db 3795 TACCGCCAG 3854  
Qy 901 cgtatgagcgag 960  
Db 3855 CGTATGGGGGGGAGGTGAGAC 3914  
Qy 961 ggtatgaaatgatttcggtgctgtcttcgaagcagtcctctagctcgcgaatctgtagc 1020  
Db 3915 GGTATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3974  
Qy 1021 atgagcgcgaaag 1080  
Db 3975 ATGAGCGCGAAACCTTTACGCTCCCGAGTTGGGGGCTCTTGGAGTACCCGAGGGA 4034  
Qy 1081 aagaatgctgtgag 1140  
Db 4035 AAGATGGTCTGTTGGAAGAGAGATTAACGATTTGCTCAACCTGCACTGGATGATGATC 4094  
Qy 1141 aaccattatcgtctgacatctgacgagcaagcgcaagcgcaagcgcaagcgcaagcgca 1200  
Db 4095 AACCATTTATCGCTCTACATTTGACGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCA 4154  
Qy 1201 agtctgactaag 1260  
Db 4155 AGTGTGATATAGGCTGGAG 4214  
Qy 1261 ggttccaagtcgctgagatgctgctgacagagagagagagagagagagagagagagag 1320  
Db 4215 GGTTCGAAGTCTGGTGCAGATGCTGCTGCAGAGAGGCGAGGAAGATAGCCCCCTCC 4274  
Qy 1321 ttcccaagaagcgagcgagagatccactacagtcagtcgtagagagagagagagagag 1380  
Db 4275 TTCCAGAGAGCGGCGGATACCTACAGTCCAGCTTGCGAGAGCTGAACGCTCTCT 4334  
Qy 1381 ggcagcgaag 1440  
Db 4335 GGCAGCAACAG 4394  
Qy 1441 ggcgtgctgctgagcgtctgcaaacatcgaagagagagagagagagagagagagag 1500  
Db 4395 GGCGTGGCTGTGAGACGCTGCAACATCGAATAAAGCTTTTATGATTTCTGCGCAAGGTA 4454  
Qy 1501 gggaaagtttgagttttgaglatctcaacgcaatagtgtaaatgacttcgtgaaag 1560  
Db 4455 GGGAAAGTTTGTAGTTTGTAGTATTTTCAACCCCAATAGTGAATGAATGACTTTCGTAACG 4514  
Qy 1561 atgtagcaataag 1620  
Db 4515 ATGTGCAATATAGCGGTAGAGCTATGAATAACAGCTGAGAGGCTGCAAAAAGCAAGCG 4574  
Qy 1621 gttgtgagcagcgaac 1680  
Db 4575 GTGTGGCGACGCAACCATCTCGGGGGGCTTAATAAAGGCTTAATTTCCGTTAATAAAG 4634  
Qy 1681 atgaatctggagcatgggttatagatctctgcaagaaatgacagagtgcttccctccatt 1740  
Db 4635 ATGATCTGGGGCATGGGCTTATAGTCTGAGAAATGACAGATGCTGTTCCCTCCATTT 4694  
Qy 1741 caaagaataatacgaagaaac 1800  
Db 4695 CAAGAATAATACCGCAAAACACCTAACAGCAAGTATATGATAGGCTGATGAACAACATG 4754  
Qy 1801 aatgacctcagaataatcagcgcatatagagctgaaagctcgagacttaacgagtgcttat 1860

Db 4755 AAATGACCTCAAGAAATTCAGCGCATTTAGAGCGTGAAGTTGCGACTTTACGCGATTTAT 4814  
Qy 1861 ctgattgcaaggag 1920  
Db 4815 CTGATGCCAGGAGAGATGCGCACAATGCGCGACATGCGCGCTTTTCATTTTCAT 4874  
Qy 1921 caccgattag 1980  
Db 4875 CACCGATAG 4934  
Qy 1981 cctggcttcag 2040  
Db 4935 CCTGGCTTCAGAGACCTTGCTTTAAACCTGAA 4969  
RESULT 6  
AX019311/c  
LOCUS AX019311 81 bp DNA Linear PAT 07-SEP-2000  
DEFINITION Sequence 1 from Patent WO9940220.  
ACCESSION AX019311  
VERSION AX019311.1 GI:10043298  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial construct.  
REFERENCE 1 (bases 1 to 81)  
AUTHORS Gauthier,J.M.  
TITLE Method of screening therapeutic agents  
JOURNALS Patent: WO 9940220-A 12-AUG-1999;  
GAUTHIER JEAN MICHEL (FR); GLAXO GROUP LTD (GB)  
FEATURES  
source  
location/Qualifiers  
1. 81  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Synthetic construct"  
BASE COUNT 36 a 27 c 18 g 0 t  
ORIGIN  
Query Match 2.3%; Score 49.2; DB 6; Length 81;  
Best Local Similarity 81.4%; Pred. No. 0.019;  
Matches 57; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
Qy 198 gctgtcgagctgtgctgagctgtgctgagctgtgctgagctgtgctgagctgtgctgagctgt 257  
Db 75 GCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 16  
Qy 258 cccgcctgtc 267  
Db 15 CTGGCTTGTCT 6  
RESULT 7  
AF429315  
LOCUS AF429315 125020 bp DNA Linear PRI 18-JAN-2002  
DEFINITION Homo sapiens junctionalilin 3 (JPH3) gene, partial cds.  
ACCESSION AF429315  
VERSION AF429315.1 GI:17646244  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 125020)  
AUTHORS Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,  
Ingwersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,  
Potter,N.T., Ross,C.A. and Margolis,R.L.  
TITLE A repeat expansion in the gene encoding junctionalilin-3 is  
associated with Huntington disease-like 2  
JOURNAL Nat. Genet. 29 (4), 377-378 (2001)  
MEDLINE 21583737  
PUBMED 11694876



QY	462	ttcaatgtaggggaatccccctaatctattcttcccatgaggaaagaacaacagtgc	521
Db	50578	tttgatgtaggggaatccccctaatctattcttcccatgaggaaagaacaacagtgc	50637
QY	522	cgaacaccggagccttcgacccaacacaacaaactgtgtcccttcgcgaagtgcgagaaggat	581
Db	50638	ycgwgcnmk-mbdwmbabmbmkgmsmdgtkdkywcvcyysmsadvybwmrbdtswgsbshy	50696
QY	582	ctatatgcgaatctcgccccgccttcgaagcgctcaagctcaatgatattatggaac	641
Db	50697	mgvhrhgvnwrgygbvcdthvyldgshmswmmkscgtsdbgkcafbkskctstmy	50756
QY	642	tgcgacgcgcgcatgtgcgcatgcatgtgcgcataatgctgcctgcgcgaatctgcg	701
Db	50757	tcgtgkykwmssasdbstyyssbshnbwsgymccrkbstvasvdrvkrkptgmyrvyrgss	50816
QY	702	aatcgacgcatgataaaacacatgacccgtgagacgctgaccccgctgttcgagaagctagc	761
Db	50817	wyyvkvbsstvasrksrkgkhyvbsmksccrrrdmstswmrsmkrcrcrshshnkskm	50878
QY	762	cctcgcgagtggtgaccccatgcatgacccctcaagaatgacgtga--cagtcggcgcttg	819
Db	50877	gkwtslvddctacsssmktkmhrkthkxkdsnhsbwksmhmkgamrvrmytsmrcsynd	50933
QY	820	gtcgaatgagcggaatagactacgcgcgaaggaacagcgcaactctctagtcagtgg	879
Db	50937	sswvdsrrrmkgsaarmgkmcycytsmsrstksrsmksmwpmgskcyycygtwscrkms	50996
QY	880	accttcggagtaactctcgtgtgtagggcgggagtcgaaccacgcggccattctcgac	939
Db	50997	mcgststggsssmkcmkycycwagrsrckscsrsrsmaggsasasmgmkmtsgnsgc	51056

## REFERENCE

REFERENCE 1 (bases 1 to 1969)







```

CDS      4009..4101
         /gene="F08G5.1"
         /product="tRNA-Arg"
         /complement(11686..11758)
         /gene="F08G5.t1"
         /note="Predicted using GeneFinder"
         /codon_start=1
         /protein_id="CAA94583.1"
         /db_xref="SPTREMBL:O19220"
         /translation="MEVNSRKLIVHGVEPIEYVCMNEVSRPSAKMKNLFIYV
         NQAQGFRLTREGEDRENTLSTYRKFYISETPEPKDLNNSNTVDWYNSVERKRGSS
         SSQPTFSQPCIDIDIGSVKSAARLSTFTSTNIGSSRMPLATNEFFSQVYMPY
         NRPASASTVSSSISGSVSVLHDDSPFSGSPSNHSLQPCPSHSSSFGSCGSS
         SHSSQSSPANAPSEPDHPSPSANDIQIDOLNPVTAPOLIKQADKCVQDROKQYD
         ERDDPAFLRRYRIKRYMSDSKMARLVGMARSEIRKLPNDEDAFYKKVVDHQRTASS
         SSKIGPAPVQOTSMEF"
         join(4263..4329,4376..4555,4855..5113,5309..5738,
         5788..5943)
         /gene="F08G5.2"
         /join(4263..4329,4376..4555,4855..5113,5309..5738,
         5788..5943)
         /note="Predicted using GeneFinder"
         /codon_start=1
         /protein_id="CAA94584.1"
         /db_xref="GI:3875593"
         /translation="MEANDKKGWLLIDNYISCLLYVAGIVAFICPIYVLCFF
         TAAPIMLSTERFRKVKIMARAGNIQFTTCRITGKYMSSGDGLERLIDCALILP
         NHLGLFDEHFEWTAADSEGVNAVGRWIFVYIMWYISPLGMSSEKSGANFTISVPHK
         RQGLHSLRRHEDRIYHEVDRLMLCLYPEGSSRLIQLKNSFEKRGKGLPFLNCAYP
         RGLAALSAIKVIGDPDPNPKLNNKNGKPLEYLDVTLGYPDGNILPLNIEFMSARN
         GCGFPFAIHEYIEFKMDPKWKDEELKNFLFEDRQKKDKLESYYTKGSPSSSEKQAF
         PSDMLIGHFILMIGLFFALVIFILSFVY"
         complement(join(6848..6998,7613..7761,7964..8053,
         8357..8473))
         /gene="F08G5.3"
         /complement(join(6848..6998,7613..7761,7964..8053,
         8357..8473))
         /note="Predicted using GeneFinder"
         /codon_start=1
         /protein_id="CAA94582.1"
         /db_xref="GI:3875591"
         /translation="MSYOEYVYKAPTEDNEHIALIRGKPSCKKANIYVNTQFOFRL
         TYLDRYDHLRKMTJSLRKVNELEIDMLNDEEMPTIOKLIDIKKQOLMRDEA
         HIKGELSLKTMNEDIRRLAGEWKKTIDARKSSAESAESSMIGDDPTSTSDGH
         KPYRETSW"
         complement(10702..10773)
         /gene="F08G5.t3"
         /note="CTG Gln O-tRNA
         Predicted using tRNAscan-SE-1.11
         Preliminary prediction
         similar to tRNA-Gln"
         /product="tRNA-Gln"
         /complement(10702..10773)
         /gene="F08G5.t3"
         /complement(11009..11080)
         /gene="F08G5.t2"
         /note="TTG Gln Q-tRNA
         Predicted using tRNAscan-SE-1.11
         Preliminary prediction
         similar to tRNA-Gln"
         /product="tRNA-Gln"
         /complement(11009..11080)
         /gene="F08G5.t2"
         /complement(11686..11758)
         /gene="F08G5.t1"
         /note="ACG Arg R-tRNA
         Predicted using tRNAscan-SE-1.11
         Preliminary prediction

```

```

gene      similar to tRNA-Arg"
         /product="tRNA-Arg"
         /complement(11686..11758)
         /gene="F08G5.t1"
         /note="Predicted using GeneFinder"
         /codon_start=1
         /protein_id="CAA94581.1"
         /db_xref="GI:3875590"
         /translation="MSAKYVVTIASCSGLAIYACLTGVAIFKINDLYDNVDDMD
         EFKMFANNAKMDIPIVTRPSLDNOSNLRAIFGEKRGAGCNGAGANNPCPGPPG
         GARGAPGDGAGAGACTGTNGISLISHRESCICPGAGEAGPAGPAGAPGPG
         QPCQDPAGCPGAPGAPGPEGDAGAPEDDAAPAPAPGAGDQGRGTCLPPAPAPGQ
         GPGNPGQDGAAGAPGQAGPAGPAGPQAGQAGQDEAGPEGNAQPGADAAVCPCP
         ARTGAVENKPEFTSGYRRRVKVVY"
         join(23649..23704,23750..23840,24635..25069,25364..25666,
         25854..26015,26074..26241,26402..26465,26511..26564,
         27016..27170,27239..27395,27512..27936,27983..28079,
         2842..28762)
         /gene="F08G5.5"
         /join(23649..23704,23750..23840,24635..25069,25364..25666,
         25854..26015,26074..26241,26402..26465,26511..26564,
         27016..27170,27239..27395,27512..27936,27983..28079,
         2842..28762)
         /note="Similarity to Rabbit UDP-glucuronosyltransferase
         (SW:UDBF_RABIT), contains similarity to Pfam domain:
         PF00059 (Lectin C-type domain), Score=15.2,
         E-value=0.00032, N=1; PF00201 (UDP-glucuronosyl and
         UDP-glucosyl transferases), Score=104.2, E-value=8.3e-29,
         N=1"
         /codon_start=1
         /protein_id="CAA94585.1"
         /db_xref="GI:3875594"
         /translation="MPYASNMGVARRPVVPMTFILSPVEYGNRPAPFLTRPRRIE
         YFTDPSQCTGFLCEFYKLSQNKQITVSGPAAVSESGKILGTSFKKTPON
         SIIITVNDPPTSEFVQDRNDLIRVGBHCYSANLSEFDFVEDSKGCTIYVAKGS
         YLNEEDGPKGTVAEFLVYEDRYVWSHKNVGAKHVSVMNAPKRELLAHEK
         TYAFITHGGLKSAKEGVCSCVPMFLPEFYDQPRMARHRTVNGIAELVYKKAITSLDI
         QOKLEKLVDPSSYKNNYKQAIKAIADISIERVYRPVPEKSPKIKISFKNVSVLY
         KNIYSIRGCGNMVYLSYLDAPISLDGAHISOVLRTEOSQKFKFRSFMNH
         LOYINIDIFLIVITRLMFVWLIFSHFPCNMHIDELFNKTKAEDITGLRDOOH
         RAYKKKKESALADASOPNLVADCAVLPCEGMHOPPTGCCYRSDASEMHGCT
         AITKALNPDAEMASFSQASIFVANKYSTHMYWLTISASAPDSASCEVMDGLL
         GLIFALSLQGNNAVSCDESIQILICKYCRETTSTVPTTKATIKTSTTVTP
         NTVYTKAPPTTKSTSTTPKFTTKITQTKKITSPTSTASIPATCTSCNCPVAPV
         NFKGICAKMCRGLVKFEDSCWNGTAAITSSGEENDEVRVSGNDEFTROIANGV
         ESSGGLNWSGQETKPDGIDYICISMDLSAGSTRKRYKILPCQSTVNSICVWNP"
         join(29612..29663,29708..29817,29871..29983,30029..30147,
         30196..30308,30695..31054,31781..31951)
         /gene="F08G5.6"
         /join(29612..29663,29708..29817,29871..29983,30029..30147,
         30196..30308,30695..31054,31781..31951)
         /gene="F08G5.6"
         /codon_start=1
         /protein_id="CAA94586.1"
         /db_xref="GI:3875595"
         /db_xref="SPTREMBL:O19223"

```

Query Match 2 1%: Score 44.6; DB 3; Length 32784;  
 Best Local Similarity 68.1%: Pred. No. 0.97; 29; Indels 0; Gaps 0;  
 Matches 62; Conservative 0; Mismatches 29; Indels 0; Gaps 0;



Oy	179	aggattactgatatgttcctgcctcgcggcgcttgcctgcgatgcctgcgcgcttgccagcttg	238
Dbl	21417	AGCATTTCTTCCTGGGTCCACGCGTCTGCCTTGATGCCATTGTCTCCTCACCTGTGCTCGTGG	21476
Oy	239	tccgccctccctcttgtcccgcgcgtacct	269
Dbl	21477	TCCGCGCTGCTCCAGCTGGTCCAGCTTGTCCT	21507
RESULT	12		
AF429315/c		125020 bp	DNA linear PRI 18-JAN-2002
LOCUS	AF429315		
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.		
ACCSSION	AF429315		
VERSION	AF429315.1	GI:17646244	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 125020)		
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L. A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2		
TITLE	Nat. genet. 29 (4), 377-378 (2001)		
JOURNAL	Medline 21583737		
MEDLINE	PUBMED 11694876		
REFERENCE	2 (bases 1 to 125020) Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L. Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA		
AUTHORS	Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA		
JOURNAL	Title Location/Qualifiers		
FEATURES	source source		
CDS	1..125020 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q24.3: between D16S520 and WI-12410" /note="Isolated from a patient with Huntington's Disease-Like 2 (HDL2)" complement(35581..35746) /rpt_type=tandem /rpt_unit=ctg complement(<36507..>36887) /gene="JPH3" /product="junctophilin 3" complement(<36507..>36887) /gene="JPH3" /note="JP3" complement(<36507..36887) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum"		
mRNA	/codon_start=1 /product="junctophilin 3" /protein_id="AAI40941.1" /db_xref="GI:17646245" /translation="MSSGRNFDPDGSSYCGWEDKRAHGVCVTPKGQGEYTGSWSN HGFEVLGYTWPSGNNTYGQTWAQRCHDIGESKKWKMYKEWMTHGFGRGVRECAAG NGAEFGMWISGLDDGYETYSNSG"		
BASE COUNT	29056 a 32731 c 30696 g 28283 t	4254 others	
ORIGIN			
Query Match	2.1% Score 44.2 DB 9 Length 125020;		
Best Local Similarity	11.3%; Pred. No.1.6; Indels 6; Gaps 1;		
Matches	98 Conservative 345 Mismatches 417 Indels 6 Gaps 1;		
Oy	393 eggsgtaatltyggccgaanaaggagttaaattgygccccgaagaagagtaatltyggcca	452	
Dbl	: : : :   :   :   : : : : : : :		

Dh	17830	MKRRMSAHSRKMMWRAYISMMRCAGMGAACAASBSRMCKCWMGIRKSRSASKSRSRGTCGYM	17771
Oy	453	tatcgcttgcttaactatgaggagaagaaaccctctaatacatttcctcccatgtggaaaagacaac	512
Dh	17770	KKKGGKGSMSKMKKKGSTSHRRSSSAKSCSYMMGMSCMSCMMAKSVMYCYYMR	17711
Oy	513	acaagtgcgcgaagacggcgcttctgcaccagacaaaacctgtctccctgcgcgaagtgc	572
Dh	17710	SSYMSVYVKCYSCOMCMSSTYSVSCCMKMSSWGSCYKCMKYSGMSYSTSMGYSTSNC	17651
Oy	573	gagaaggcttatcatgccaalctgcgcccgccctgcgaagcgctcaagcatcatgatattat	632
Dh	17650	KYKKSMSMSMYCKTISKSYRKRRTSYIYWGCGRAKAKKYYCAGRRTMRMYWKCAMWMSY	17591
Oy	633	galaaccaactgcggcgccgcgcgatgctgatgatgtgcgcgaatgaatgatgcgctlgccga	692
Dh	17590	CCMSYCMTHYYYSKSTCTYKSSFTCYKRGGYWGSKTICYSAGKSRSMYCYMHSSSKSSWSM	17531
Oy	693	cattgcgcgaatcgacgcgcgatgaaaaaaccaatgacgcgtgaagaccctgcaccgcttgtga	752
Dh	17530	SMARSSMCMGWAGYRKRKRSAGWAGMRSSSGKSRSTMKRCASKTSYSTGTSMKK	17471
Oy	753	ggaagctacccgttggygtgtttaccatgatgatgacctcaagaatgacgttgacaagctgc	812
Dh	17470	KCYSVYSNGMKKKATTCMCMTKYRYRTSMCMWTYKMSWGYKTKRKCCMKKKGCTGYRG	17411
Oy	813	cggcttgctgcatagtgagcgcgatatagatacccgcagaagcggaagcggaactcctagt	872
Dh	17410	MSSKKSQYKSMHGMHGSSYSTSCMKSCMGYSMMKCMKYMSYKKRRSRMRSGMSKGMG	17351
Oy	873	gaecgtgaaacttcgcgaatatacttcgcgtcgtatgtgcggcggaatcgaaaccaactggccat	932
Dh	17350	AGRCGYSSSMWSTRKRRSKCYKSKYGKGRKMGCMKRGSKYWSSMKMKMRSSWSGC	17291
Oy	933	tctgcacgcctcaaaagcatlctcacctctg-----gttagtaagtcttcgcgtcgtctt	986
Dh	17290	YSTKYIKGGRKSKGWGSTKSKAASSNMHMASKCTYSSSYWSNNRRNRNMGKTCNYM	17231
Oy	987	ccagcagctctctaatcttcgcgaactctgatactgcgaatgagcggaanaaccttaagctcc	1046
Dh	17230	RISRRAAMGNNAAACCTTCCCANTNNGGGGAAAAAGGCCGASASASCYKGMSSKCYR	17171
Oy	1047	cgaagtcggcgcgctctcttggagtgcccgagggaaagaatgylgtctgttggaacgacgttaa	1106
Dh	17170	SGTRRCMKSKCRSGRSRGKSMGMTGRSGGKTSYSAKSGRCCYCMCWGRKRCMSSR	17111
Oy	1107	cagatttgctctcaaacctgcacgtgatgagatacaacattatctgcgtctgtaacttac	1166
Dh	17110	KMKMKYYRRRRRWMTCMKCSMYMAMYCRMCMCMCKSCCGCYSMGMSSTYSYSGYS	17051
Oy	1167	ggcaagccgcgaacagatgtgcgcgtgcgttgcaagctgtgatactgaagctgggaagtga	1226
Dh	17050	WCMKSYMRSYTSKRSSTSLAMRSSKRGMDTGTRYGGGRSYGKGSGYSTKGGSWGGKGGK	16991
Oy	1227	agacgaccacacgctgcgcgaagcgcg	1252
Dh	16990	STRGSSAGKSSYKKMKSCCARFYKSMS	16965
RESULT_13			
BTA414556			
LOCUS	BTA414556	3490 bp	mRNA linear MAM 05-OCT-2001
DEFINITION	Bos taurus mRNA for Ikb kinase-beta (bikbeta gene).		
ACCESSION	AJ141456		
VERSION	AJ141456.1	GI:15986410	
KEYWORDS	bikbeta gene; Ikb kinase-beta.		
SOURCE	COW.		
ORGANISM	Bos taurus		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
Bovidae; Bovinae; Bos.			
1 (bases 1 to 3490)			
REFERENCE			
AUTHORS	Rotenberg,S., Dobbelaere,D.A.E. and Heusler,V.T.		



TITLE	Identification and characterisation of the bovine Ikb kinases (IKKs) alpha, beta and gamma
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3490)
AUTHORS	Rottenberg,S.
TITLE	Direct Submission
JOURNAL	Submitted (12-SEP-2001) Rottenberg S., Molekulare Pathologie, Institut fuer Tierpathologie/Uni Bern, Laengassstrasse 122, CH-3012 Bern, SWITZERLAND
FEATURES	Location/Qualifiers
source	1..3490
	/organism="Bos taurus"
	/db_xref="taxon:9913"
	/clone="plate 77, row E, column 16; USDA, ARS, US Meat Animal Research Center"
	/clone_1lb="MARC 4BOV"
	/tissue_1lb="pooled tissue from day 20 and day 40 embryos"
	36..2306
	/gene="bikKbeta"
	36..2306
	/gene="bikKbeta"
	/function="Signalling molecule of the NFkB pathway"
	/codon_start=1
	/product="IKB kinase-beta"
	/protein_id="CAC93687.1"
	/db_xref="GI:15986411"
	/translation="MSMSPSLPTQTCGAWEMKRLGTGGGVIRHNOETGEQIAIK QCRQLSPNRERWLEIOTMRLNHPNVARADPEQMSIAPNDILPLAEMCOGG DLKRYLYNFGCCGREGAIIITLDSIALRLEENRIIHRDKRENTVLDQGBRL IRIIDIGAKELDQSLCTSVGLQYLAPELLEQKTYLVVDWSPFTLAQECTG FRPFLPMQPVOMHSKVRKSEMDIVASDLNGAVFSSLPHPNNLSVLAQRLEKW LQMLMHPROKTDIPVYGNPCFKALDILNKLILHVNVTGTLATVPVDESLQ SIKARIKDTGILEEDDELLOEGALIDPKPAOCLISGKLINEGRITLMDIVFLDN SRKYTSQVSPQSPESVSCITQEPKRNLPFQLRKVMQVMSHIOALKEQCSRILOG QRAAMNLLRNNSCLSKRNKSNASMSQKATLDFPKTSIOTIDLEKRYROTEGTLSD QRLAWREMQAVELCGRENEVNHVERMALDITDVLQDSPMGRKQGTDLDDLEQ ARLELYRRLREKPDQTDSDQEMVRLLOAIQGFKVRVITYTQLSKTVCKQKLE LIPKVEEVSIMSDEKMYVRLQEKROKELMNLKTAQSKVRGVPVSGSDSNASHL S HPCOLMSOCTAPDSLPEAAEKSESDLVAEHNLCTOLENALQDTMKEQDOSLSRLDMS WLOSREEOOSLERAS"
BASE COUNT	779 a 1024 c 1064 g 623 t
ORIGIN	
Query Match 2.0%: Score 43.2; DB 4; Length 3490;	
Best Local Similarity 48.4%; Pred. No. 1.7;	
Matches 120; Conservative 0; Mismatches 128; Indels 0; Gaps 0;	
Qy 606	gcaaggcgctcaagctcatgcatgattaatgataagccactgagggcgccgcatgctgata 665
Db 1730	gcagggcgggagcgtttacccagcgctgagggagagagcccgagaccgagctgagg 1789
Qy 666	tgtagccatgaatgtagcgctgagccagatcgcccaatgacgcatgataaacatga 725
Db 1790	tgacagtcaggaatgtagcgctgagccagatcgcccaatgacgcatgataaacatga 1849
Qy 726	ccgtgagagcctgagcccgctgttcgagagagtagcgctgaggtgttgcacatgata 785
Db 1850	gcgagtgattttatgacgagctcagtaagaccgctgtgcaagcgaagagccctggaact 1909
Qy 786	ccctgcaaatgatacgtagaacatcgagcgctgtgtgtagatgagcgcaataagactacg 845
Db 1910	gctgcccacgagtgagaggtgtgtagcctgattgagcagagcgaagagatgtgtgctcg 1969
Qy 846	ccagagag 853
Db 1970	gctgcagc 1977

RESULT 14  
AE005932 13291 bp DNA linear BCT 28-MAR-2001  
LOCUS AE005932 section 258 of 359 of the complete genome.  
DEFINITION Caulobacter crescentus

ACCESSION	AE005932 AE005673
VERSION	AE005932.1 GI:13424216
KEYWORDS	
SOURCE	Caulobacter crescentus.
ORGANISM	Caulobacter crescentus
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter.
AUTHORS	1 (bases 1 to 13291)
	Nierman,W.C., Feldblyum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Pirocka,I., Nelson,M.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Debroy,R.T., Dodson,R.J., Durkin,A.S., Gwin,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Knouti,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Yamahayan,J., Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and Fraser,C.M.
TITLE	Complete genome sequence of Caulobacter crescentus
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
MEDLINE	2 (bases 1 to 13291)
REFERENCE	Nierman,W.C., Feldblyum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Pirocka,I., Nelson,M.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Debroy,R.T., Dodson,R.J., Durkin,A.S., Gwin,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Knouti,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Yamahayan,J., Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C. and Fraser,C.M.
AUTHORS	Direct Submission Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
source	1..13291
	/organism="Caulobacter crescentus"
	/db_xref="taxon:69394"
	79..3009
	/gene="CC2638"
	79..3009
	/gene="CC2638"
	/note="Identified by match to PRAM protein family HMM PF00675"
	/codon_start=1
	/transl_table=11
	/product="peptidase, M16 family"
	/protein_id="AAK24605.1"
	/db_xref="GI:13424217"
	/translation="MISASRLAIVAAAGLSVAAACSQLPSPSIPLLKKDAGQASAKTP DAPALAGPVYTDLPQMAOELSDVADPDPAWGRGVLPNOMRALRNKNAPPCQAALRL MIDAGSMSEADDDQGLAHLEHMAFNKSNVTEGEMIKILIERGLAFGADTNAISFD ERTYQIDLRKTDQDDTVDTSIMLREAAAGELLIPAEVADREGRVLSSEERTPTGVRV AITLSNOMEGOLPKRPIPTGKREYVKTAPQAIRDFEYEAIVPERSIVLAAGDFVD LEMFAKTIGKFGDWGKGNPKDPDVPYAKRGTAPMPEAGAPMSIOMTWRRKPGGL LKTRKAVDERDTLENGFPAVINRRLQAVGSAEPTTAGCAFKDQDQCAIVRTTGGTA QPGRMBEALITADAEORRAILOYGVRODELDRELTAJLRAGIIVAAAGEAIVORPSLANQ LVGTLDGGEVYTPESONLAFDLFVKGGLAEIRVNAIYKSAFVSGGLVLAAPTAVNG GEPALIKAYAEELKAQPAPIAGVYVPEYSSPGKGAQAEQVEVDLVAIVRPNVG VRLTVKPTKFRDQOVIVYKVRAGGLDMPDSKOSPLMSGAFIEGILKQITQDMERY LKTRKINAOLGVGDDAFETLNGRTPREDLISRELOVLAAPTEVREAREAFNRKISYGT LHQLOSTTGKAVGIDGILMBGSDQRMTPSPSEQJASISDILKAAVNPILAKGGLLE VVIVGDTTVDKAIAAVADTFEGALPARGDDPLGAEKAPPAPASKEPVYIRTHNGRPDO AALFMVTRTDLDEPSNLORSRDVSLQVOMLRLTIDELREKQATYSPLNATASVYFD DMGVLAVSLDEPPEKLDGVVASIRKIIAADLRDPSADEIDRAKRRKIDAIEARVTN EYVWVGALSGAHTDPRLLDARTSVIAGISRVTPADVOKAAQTYIADERSWLLVKKPRAA AAK"
	complement(3071..4000)
	/gene="CC2639"
	complement(3071..4000)
	/gene="CC2639"
	/note="Identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAK24606.1"

gene  
CDS



[illegible][illegible]







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 15:36:33 ; Search time 242.01 Seconds  
(without alignments)  
14983.361 Million cell updates/sec

Title: US-09-826-206-1

Perfect score: 2112

Sequence: 1 ggcattggctcgaattcatc.....tgaggaaataaggagacgtg 2112

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: N.Geneseq\_032802.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.\*  
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.\*  
25: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2112	100.0	2112	24	AA518306
2	2112	100.0	2112	24	AA518308
3	2007	95.0	2517	24	AA518309
4	2007	95.0	2517	24	AA518309
5	2007	95.0	2517	24	AA518309
6	2007	95.0	2517	24	AA518309
7	40.4	1.9	731	13	AA021036
8	40.4	1.9	1980	23	AA52512
9	39.8	1.9	10886	24	AA581796
					AB134134

C 10	38.8	1.8	4590	22	AAH24065	Yeast A09604-asso
C 11	38.2	1.8	2047	23	AA587283	DNA encoding novel
C 12	37.4	1.8	1456	24	AA517123	Replicon sequence
C 13	37.4	1.8	1458	24	AA517226	Replicon sequence
C 14	37.4	1.8	7029	24	AA517119	Ketogulonigenium e
C 15	36.8	1.7	858	23	AA554178	Pseudomonas aerugi
C 16	36.6	1.7	36471	21	AAAB1453	N. meningitidis pa
C 17	36.6	1.7	349980	21	AAAF21611	Neisseria meningit
C 18	36.4	1.7	2035	23	AB110875	Drosophila melanog
C 19	36.4	1.7	4146	23	AB110875	Drosophila melanog
C 20	36.2	1.7	1310	21	AA548039	Zea mays DNA fragm
C 21	35.8	1.7	4403765	22	AA199683	Mycobacterium tube
C 22	35.8	1.7	4411529	22	AA199682	Drosophila melanog
C 23	35.6	1.7	6692	23	AB119806	Human OREF ORF103
C 24	35.4	1.7	926	21	AA57548	Human CDNA sequenc
C 25	35.4	1.7	1982	22	AAH15840	Human digestive sy
C 26	35.4	1.7	3046	21	AA589786	Gene encoding a su
C 27	35.4	1.7	10732	21	AA510594	Gene encoding a su
C 28	35.4	1.7	10732	21	AA510594	Drosophila melanog
C 29	35	1.7	642	23	AB121677	Drosophila melanog
C 30	35	1.7	2642	23	AB121676	Wheat glutenin gen
C 31	34.8	1.6	1302	24	AA170852	Pseudomonas aerugi
C 32	34.8	1.6	2178	23	AA51488	DNA encoding human
C 33	34.6	1.6	3174	21	AA529843	Human immune/haema
C 34	34.6	1.6	17687	22	AA564966	Human immune/haema
C 35	34.6	1.6	17687	22	AA564966	S. spinoza DNA fra
C 36	34.6	1.6	45624	22	AA588315	S. spinoza DNA fra
C 37	34.6	1.6	50000	22	AA588312	Human cDNA clone (
C 38	34.4	1.6	586	22	AAH10634	Drosophila melanog
C 39	34.4	1.6	2757	23	AB112845	Drosophila melanog
C 40	34.4	1.6	2966	23	AB112843	Human silt-3 prote
C 41	34.4	1.6	5015	20	AA589163	Human silt-3 encod
C 42	34.4	1.6	5015	20	AA589163	Human silt coding
C 43	34.4	1.6	5176	22	AA51026	Quorum sensing con
C 44	34.4	1.6	7110	20	AA51351	Human immune syste
C 45	34.4	1.6	7449	24	AB132277	Human immune syste

## ALIGNMENTS

RESULT 1	
AA518306	AA518306 standard; DNA: 2112 BP.
ID	AA518306
XX	AA518306:
AC	AA518306:
XX	12-MAR-2002 (first entry)
DT	12-MAR-2002 (first entry)
XX	DNA sequence of Ketogulonigenium replicon from plasmid PADM291.
DE	DNA sequence of Ketogulonigenium replicon from plasmid PADM291.
XX	Cloning vector; Ketogulonigenium replicon; endogenous plasmid;
KW	transformed host cell; Escherichia coli; PADM291; ds.
KW	transformed host cell; Escherichia coli; PADM291; ds.
XX	Ketogulonigenium sp. strain ADM291-19.
OS	Ketogulonigenium sp. strain ADM291-19.
XX	WO200177347-A2.
PN	WO200177347-A2.
XX	18-OCT-2001.
XX	18-OCT-2001.
PD	05-APR-2001; 2001WO-US11059.
PF	05-APR-2001; 2001WO-US11059.
XX	05-APR-2001; 2000US-194625P.
PR	05-APR-2001; 2000US-194625P.
XX	(ARCH ) ARCHER-DANIELS MIDLAND CO.
PA	(DELTA ) ARCHER-DANIELS MIDLAND CO.
PA	(DELTA ) ARCHER-DANIELS MIDLAND CO.
XX	D'Elia J.
PI	D'Elia J.
XX	WPI: 2002-049150/06.
DR	Novel nucleic acid vector comprising Ketogulonigenium replicon found on
XX	a specific deposited endogenous plasmid, useful for producing
PT	a specific deposited endogenous plasmid, useful for producing



polypeptides and/or transcripts by culturing host cells transformed with vector

Claim 2: Fig 1; 66bp; English.

The present invention relates to the isolation of vectors comprising a Ketoglulonigenium replicon found on the endogenous plasmid, pADM291. The invention also describes methods of transforming host cells with the vectors and producing polypeptides and/or antisense transcripts by culturing the transformed host cells. The vectors are useful for transforming a host cell by conjugation or electroporation. The vectors which have a replicon functional in both Ketoglulonigenium and Escherichia coli, enable the cloning of certain genes of Ketoglulonigenium in E. coli as the latter is an efficient host for amplification of vector DNA. The present DNA sequence represents the replicon of Ketoglulonigenium endogenous plasmid pADM291.

Sequence 2112 BP; 530 A; 509 C; 605 G; 468 T; 0 other;

Query Match 100.0%; Score 2112; DB 24; Length 2112;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ggcgaatggtcgaataatcatagaattttgtgtgaggtgctgaagcgtctcgacaaggtgtg 60  
D 1 ggcgaatggtcgaataatcatagaattttgtgtgaggtgctgaagcgtctcgacaaggtgtg 60  
Y 61 ctgcgcgagatctctgtgtcctcagtagggcgacaatgagaggtgttaattgccccctg 120  
D 61 ctgcgcgagatctctgtgtcctcagtagggcgacaatgagaggtgttaattgccccctg 120  
Y 121 tatcgtctcgcgtggtgcatttgggtatctcgtccgacatatgatatcgctagag 180  
D 121 tatcgtctcgcgtggtgcatttgggtatctcgtccgacatatgatatcgctagag 180  
Y 181 gattactatagttctcgtcgtcggtgttcgagcttgcggagctgtgcggctgtgc 240  
D 181 gattactatagttctcgtcgtcggtgttcgagcttgcggagctgtgcggctgtgc 240  
Y 241 gggcgtctccctctgtccgcgtgttcctcaattttcaacaatcaaaaatggcgaaagc 300  
D 241 gggcgtctccctctgtccgcgtgttcctcaattttcaacaatcaaaaatggcgaaagc 300  
Y 301 cctcttctcatatagttcttatagttatacgaataatcacataatcatatagctt 360  
D 301 cctcttctcatatagttcttatagttatacgaataatcacataatcatatagctt 360  
Y 361 attcgtcttaaaaggagtaattggtccgacaagaggaatattggcgccgcaaaaggagtt 420  
D 361 attcgtcttaaaaggagtaattggtccgacaagaggaatattggcgccgcaaaaggagtt 420  
Y 421 aattggcgccgcaaaaggagtaattggtccgacatcggtgtttataatggaggaatcc 480  
D 421 aattggcgccgcaaaaggagtaattggtccgacatcggtgtttataatggaggaatcc 480  
Y 481 ccttaatcatcttcccatggaagagaacaacaagtggtccgacagccgggcttcgac 540  
D 481 ccttaatcatcttcccatggaagagaacaacaagtggtccgacagccgggcttcgac 540  
Y 541 cagacaaaaactgtgtccctgcgaggtgtgcagaggggtctatataagccaatccgccc 600  
D 541 cagacaaaaactgtgtccctgcgaggtgtgcagaggggtctatataagccaatccgccc 600  
Y 601 cgctcgagggcgctcaagctcatgcatattatagtagccactcggtcggtccgcatgtgt 660  
D 601 cgctcgagggcgctcaagctcatgcatattatagtagccactcggtcggtccgcatgtgt 660  
Y 661 gatgatgtgcgcataagaatgcgtgtccgacacatttcgcaatcgaacgcatgaaaac 720  
D 661 gatgatgtgcgcataagaatgcgtgtccgacacatttcgcaatcgaacgcatgaaaac 720  
Y 721 catgaacgtgtagacccgacccgctgtcgaagagctagcgtgtggtttgaccat 780

Db 721 catgaacgtgtagacccgacccgctgtcgaagagctagccgctgtggttaccat 780  
Y 781 gatgacctgcaagatgatctgtacagctgcgtgtgtgtcatgtgagcgcaatagac 840  
D 781 gatgacctgcaagatgatctgtacagctgcgtgtgtgtcatgtgagcgcaatagac 840  
Y 841 taccgcagagagcaagcggtgaactctctagtcggtgagaccttcgagttacatccgt 900  
D 841 taccgcagagagcaagcggtgaactctctagtcggtgagaccttcgagttacatccgt 900  
Y 901 cgtatggcggtgaatcgaacacttggtgacatctcgaacgctcaaacggtatctctc 960  
D 901 cgtatggcggtgaatcgaacacttggtgacatctcgaacgctcaaacggtatctctc 960  
Y 961 gtagtaaglatccgtgtctgtcttcagacaggtcttacttcgccaacttgatcg 1020  
D 961 gtagtaaglatccgtgtctgtcttcagacaggtcttacttcgccaacttgatcg 1020  
Y 1021 atgagcgcaaaaccttaacgtgtcccgagttgcgtggtctccttgaggtccgagga 1080  
D 1021 atgagcgcaaaaccttaacgtgtcccgagttgcgtggtctccttgaggtccgagga 1080  
Y 1081 aagatgtgtcgttggaacgacgttaacagatttgtctcaaacctgcacatgtagatc 1140  
D 1081 aagatgtgtcgttggaacgacgttaacagatttgtctcaaacctgcacatgtagatc 1140  
Y 1141 aaccatttctcgtctgtgacatgtacgcaaacgcaacgaatgtgcgttagcgtgca 1200  
D 1141 aaccatttctcgtctgtgacatgtacgcaaacgcaacgaatgtgcgttagcgtgca 1200  
Y 1201 agtgtactataaggtcgtggagatgaaagacgacccaacgtgtgcagagcgagctgtgc 1260  
D 1201 agtgtactataaggtcgtggagatgaaagacgacccaacgtgtgcagagcgagctgtgc 1260  
Y 1261 ggttccaaggtcgttcgagatgtctcgtcgaagggcgacgcaaaagctagcccttc 1320  
D 1261 ggttccaaggtcgttcgagatgtctcgtcgaagggcgacgcaaaagctagcccttc 1320  
Y 1321 ttcccaagaacggtgcggtatcaactacatcagctgtgtgtgaggttgaaacgtctgtc 1380  
D 1321 ttcccaagaacggtgcggtatcaactacatcagctgtgtgtgaggttgaaacgtctgtc 1380  
Y 1381 ggcgcgaacgaagcaacgactctgactcgtcactcagacttcggcggttctcgtcgagaga 1440  
D 1381 ggcgcgaacgaagcaacgactctgactcgtcactcagacttcggcggttctcgtcgagaga 1440  
Y 1441 ggcgtgcgtcgtgacgtgcgaacatcgaacaaactgttttaagattcttcgcaaaagta 1500  
D 1441 ggcgtgcgtcgtgacgtgcgaacatcgaacaaactgttttaagattcttcgcaaaagta 1500  
Y 1501 gggagagtttgaagtttgaagttatttcccgcaatagtttaaatgacttcgtgaaagc 1560  
D 1501 gggagagtttgaagtttgaagttatttcccgcaatagtttaaatgacttcgtgaaagc 1560  
Y 1561 atgtcaatataagcgttaagaactatgaatacaacggtcgtgacaggtcgcaaaacgag 1620  
D 1561 atgtcaatataagcgttaagaactatgaatacaacggtcgtgacaggtcgcaaaacgag 1620  
Y 1621 gttgtgcgacgcgaacccatcactcgtgcgtcaaaaagcgttaaaatttcggtlaaaaag 1680  
D 1621 gttgtgcgacgcgaacccatcactcgtgcgtcaaaaagcgttaaaatttcggtlaaaaag 1680  
Y 1681 atgaatcctggagcaggtgtatagatctcgcagaaatgcagagaggtttctctccatct 1740  
D 1681 atgaatcctggagcaggtgtatagatctcgcagaaatgcagagaggtttctctccatct 1740  
Y 1741 caaagaataacacgcgaataacttaacacgcaaggtataggttaagcgttatgaaacacatg 1800  
D 1741 caaagaataacacgcgaataacttaacacgcaaggtataggttaagcgttatgaaacacatg 1800  
Y 1801 aaatgaccccaagaataacgacatlagagcgtgaaagttcgacattacgagatgttat 1860







```

|||||
Db 1362 ggtagtaagatcctcgtctgtctgttccagacgctctagctcgcacatcttgatcgg 1421
QY 1021 atgagcgcaaaaaacttaacggtcccgaggttcggtggtctcttggtggtcccgagga 1080
Db 1422 atgagcgcaaaaaacttaacggtcccgaggttcggtggtctcttggtggtcccgagga 1481
QY 1081 aagatggttcgttgagacgagcttaacagatctgtctcaaacctgactgtagatc 1140
Db 1482 aagatggttcgttgagacgagcttaacagatctgtctcaaacctgactgtagatc 1541
QY 1141 aacattatcgcgtctgacatctgacgagcaagccgacccaagaattggccgttagcgtgga 1200
Db 1542 aacattatcgcgtctgacatctgacgagcaagccgacccaagaattggccgttagcgtgga 1601
QY 1201 agtgtgactaagcttggtgaggtgaagagacacccaacgctgcgcagcgcgagctggcg 1260
Db 1602 agtgtgactaagcttggtgaggtgaagagacacccaacgctgcgcagcgcgagctggcg 1661
QY 1261 ggtccaaagtcggtcagagatgctcgttcgcagagggcgacgagaaacgatatagcccttc 1320
Db 1662 ggtccaaagtcggtcagagatgctcgttcgcagagggcgacgagaaacgatatagcccttc 1721
QY 1321 ttcccaagaagcggtgagatcacctacagcttggtcgtgagctgaaagctctgct 1380
Db 1722 ttcccaagaagcggtgagatcacctacagcttggtcgtgagctgaaagctctgct 1781
QY 1381 ggcagcaacaagaagacacgactgactgcctcagacattccggcgtttctgctcgagagaga 1440
Db 1782 ggcagcaacaagaagacacgactgactgcctcagacattccggcgtttctgctcgagagaga 1841
QY 1441 ggcgtgctgctgagcgtcgaacatcgaaacatcgattttagatcttctgcgcaagta 1500
Db 1842 ggcgtgctgctgagcgtcgaacatcgaaacatcgattttagatcttctgcgcaagta 1901
QY 1501 gggaaaggttgagtttgaggtatttcaccgcgaataagtgtaattgacttctgtaaaag 1560
Db 1902 gggaaaggttgagtttgaggtatttcaccgcgaataagtgtaattgacttctgtaaaag 1961
QY 1561 atgtgcaatatacggttaagaatacgaacacgctgagacagctgcaaaagcaagcg 1620
Db 1962 atgtgcaatatacggttaagaatacgaacacgctgagacagctgcaaaagcaagcg 2021
QY 1621 gttggtgacccgcaacatcatcactcgggcttaaaagcgttaaatcttcgggttaaaag 1680
Db 2022 gttggtgacccgcaacatcatcactcgggcttaaaagcgttaaatcttcgggttaaaag 2081
QY 1681 atgaatctgggtgcatgtgataatcctcgaagatgcaagaatgttctccatctt 1740
Db 2082 atgaatctgggtgcatgtgataatcctcgaagatgcaagaatgttctccatctt 2141
QY 1741 caaagaataacacgcaaacacactaacaacgcaatatagtgtaaacggtgagaaacacatg 1800
Db 2142 caaagaataacacgcaaacacactaacaacgcaatatagtgtaaacggtgagaaacacatg 2201
QY 1801 aatgaactcagaatacgcgcgcatatagacgttgaagtgcgaacttaccgcatcttat 1860
Db 2202 aatgaactcagaatacgcgcgcatatagacgttgaagtgcgaacttaccgcatcttat 2261
QY 1861 ctgatatcgagagagatcgcgacaataatggcgagacatggccgaagcttccaattcat 1920
Db 2262 ctgatatcgagagagatcgcgacaataatggcgagacatggccgaagcttccaattcat 2321
QY 1921 caacgttgagagagagacccgccccctcaaaaacaaagatggtggaagatatactctat 1980
Db 2322 caacgttgagagagagacccgccccctcaaaaacaaagatggtggaagatatactctat 2381
QY 1981 cctgggcttcaaggaccttgcttactcgtgagcaaaaacgcatatttgagacagggccc 2040
Db 2382 cctgggcttcaaggaccttgcttactcgtgagcaaaaacgcatatttgagacagggccc 2441
QY 2041 gcaacttaaggcggaagcctataacgagtaacaaaacatagaaacagatagagaa 2100

```

```

Db 2442 gcaacttaaggcggaagcctataacgagttaccaaaacactagaagccaagattgagaa 2501
QY 2101 atagggagcgtg 2112
Db 2502 atagggagcgtg 2513

RESULT 3
AAS18309
ID AAS18309 standard; DNA; 2517 BP.
AC AAS18309;
DT 12-MAR-2002 (first entry)
DE DNA region of Ketogulonigenium plasmid pADM291 supporting replication.
EX
KW Cloning vector; Ketogulonigenium replicon; endogenous plasmid; pADM291;
transformed host cell; Escherichia coli; plasmid vector replication; ds-
XX Ketogulonigenium sp. strain ADM291-19.
OS
PN WO200177347-A2.
PD 18-OCT-2001.
PF 05-APR-2001; 2001WO-US11059.
PR 05-APR-2000; 2000US-194625P.
PA (ARCH ) ARCHER-DANIELS MIDLAND CO.
PI (DELI/) D'ELIA J.
D'Elia J;
WPI; 2002-049150/06.
XX
DR Novel nucleic acid vector comprising Ketogulonigenium replicon found on
XX a specific deposited endogenous plasmid, useful for producing
XX PT polypeptides and/or transcripts by culturing host cells transformed
XX with vector
XX
PS Claim 4; Fig 4; 66pp; English.
XX
CC The present invention relates to the isolation of vectors comprising
CC a Ketogulonigenium replicon found on the endogenous plasmid, pADM291.
CC The invention also describes methods of transforming host cells with
CC the vectors and producing polypeptides and/or antisense transcripts by
CC culturing the transformed host cells. The vectors are useful for
CC transforming a host cell by conjugation or electroporation.
CC The vectors which have a replicon functional in both Ketogulonigenium
CC and Escherichia coli, enable the cloning of certain genes of
CC Ketogulonigenium in E.coli as the latter is an efficient host for
CC amplification of vector DNA. The present DNA sequence represents the
CC region of Ketogulonigenium endogenous plasmid pADM291 that supports
CC plasmid vector replication.
XX
SQ Sequence 2517 BP; 657 A; 589 C; 683 G; 588 T; 0 other;

Query Match 95.0%; Score 2007; DB 24; Length 2517;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ggcacatggctcgaataatcatagaatttggtagaggtgctgagcgctctgcacaggggtg 60
Db 1 ggcacatggctcgaataatcatagaatttggtagaggtgctgagcgctctgcacaggggtg 60
QY 61 ctgagcggaagatctctggtcaggtagggcgacaatggagaggtgttggtgccccctg 120
Db 61 ctgagcggaagatctctggtcaggtagggcgacaatggagaggtgttggtgccccctg 120
QY 121 tatcgctctcgtggtgagcattggtcatccctgcgcggacatatatattccgctagag 180

```



|||||  
Db 121 tatcgctctcgctgycgcatctggtcatctcgcgcgacatatgatctccgtagag 180  
Qy 181 gattacagatagttctgcgcgttcgggtctgcggtctgcgcggtctgtcggtctc 240  
Db 181 gattacagatagttctgcgcgttcgggtctgcggtctgtcggtctgtcggtctc 240  
Qy 241 gggcctctccctctgtccgcgcgtctccactctttcacatcaaaaacggcggaac 300  
Db 241 gggcctctccctctgtccgcgcgtctccactctttcacatcaaaaacggcggaac 300  
Qy 301 cctctctgtctatagttctatagttcatacgaaaattacataatattacatagctt 360  
Db 301 cctctctgtctatagttctatagttcatacgaaaattacataatattacatagctt 360  
Qy 361 attcgcttaaaaggagtaattggccgcaaaaggagtaattggccgcaaaaggaggt 420  
Db 361 attcgcttaaaaggagtaattggccgcaaaaggagtaattggccgcaaaaggaggt 420  
Qy 421 aattgggccgcaaaaggagtaattggccgcatatcggtctgttacaatgggaagatcc 480  
Db 421 aattgggccgcaaaaggagtaattggccgcatatcggtctgttacaatgggaagatcc 480  
Qy 481 ccttaatacttctcccatatgggaagaacaacaagtggccgcgacccggccttcgac 540  
Db 481 ccttaatacttctcccatatgggaagaacaacaagtggccgcgacccggccttcgac 540  
Qy 541 cagacaaaatactgtgtctccgcgcgaggtggcgagagggtctatattcgcaatccgccc 600  
Db 541 cagacaaaatactgtgtctccgcgcgaggtggcgagagggtctatattcgcaatccgccc 600  
Qy 601 cgactgcagcgctcaagctcatatgataatgatacgcaatcggcgcgacatggtct 660  
Db 601 cgactgcagcgctcaagctcatatgataatgatacgcaatcggcgcgacatggtct 660  
Qy 661 gatgatgtgcgcataatgaatgctggcgcgacatcgcgcaatcgcgcgcatgaaac 720  
Db 661 gatgatgtgcgcataatgaatgctggcgcgacatcgcgcaatcgcgcgcatgaaac 720  
Qy 721 catgacggtgagagccttgaccccgctgttgcagagcctacgctgcggtcttgaccat 780  
Db 721 catgacggtgagagccttgaccccgctgttgcagagcctacgctgcggtcttgaccat 780  
Qy 781 gatgacctgcaaaagatgatcatgatacagctgcgcgtctgcgatacgcgcaataagac 840  
Db 781 gatgacctgcaaaagatgatcatgatacagctgcgcgtctgcgatacgcgcaataagac 840  
Qy 841 taaccgcagagcgcaagcgcgcaactcctagtgacgtgagaccttcgcgagatactcgt 900  
Db 841 taaccgcagagcgcaagcgcgcaactcctagtgacgtgagaccttcgcgagatactcgt 900  
Qy 901 cgtatgctgcgagagctgcgaaccacttgcgcatcttcgcgacggtatctcatctc 960  
Db 901 cgtatgctgcgagagctgcgaaccacttgcgcatcttcgcgacggtatctcatctc 960  
Qy 961 gtagtaagtaattcgttgtctgtcttcacagacgtctctatctcgcgaactcttgatcgg 1020  
Db 961 gtagtaagtaattcgttgtctgtcttcacagacgtctctatctcgcgaactcttgatcgg 1020  
Qy 1021 atgagcgcgaaacaccttaagctgtcccgagttgcggcgctccttgagtgcccgagga 1080  
Db 1021 atgagcgcgaaacaccttaagctgtcccgagttgcggcgctccttgagtgcccgagga 1080  
Qy 1081 aagatgtgtctgtggaagcagtaacagattgtctccaacctgcgacgtgatgagtc 1140  
Db 1081 aagatgtgtctgtggaagcagtaacagattgtctccaacctgcgacgtgatgagtc 1140  
Qy 1141 aacctattatcgctctgacattgacgcgcaaacgcgaccaagattgacgtagcgtagca 1200  
Db 1141 aacctattatcgctctgacattgacgcgcaaacgcgaccaagattgacgtagcgtagca 1200  
Qy 1201 agtgtgactatagcctgggaagtgaagaagaccacacgctgcgcaaggcgagctggcg 1260  
|||||

Db 1201 agtgtgactatagcctgggaagtgaagaagaccacacgctgcgcaaggcgagctggcg 1260  
Qy 1261 ggttccaaggttcgggtcgaagatctgtctgcgaagaaggcgacggaaagataacccctcc 1320  
Db 1261 ggttccaaggttcgggtcgaagatctgtctgcgaagaaggcgacggaaagataacccctcc 1320  
Qy 1321 ttcccaagcgggcgagatcacctcacagttccacgtctgctgagatcgtaaacgctctgt 1380  
Db 1321 ttcccaagcgggcgagatcacctcacagttccacagttctgagatcgtaaacgctctgt 1380  
Qy 1361 ggcagcaaacaggacacgattctgacgtcctcagaacttcctcggtcttcgtcggagaga 1440  
Db 1361 ggcagcaaacaggacacgattctgacgtcctcagaacttcctcggtcttcgtcggagaga 1440  
Qy 1441 ggcgtgcctctggaagcgtgcgaacactcgaaaaacgttttatgattcttcgcgcaagta 1500  
Db 1441 ggcgtgcctctggaagcgtgcgaacactcgaaaaacgttttatgattcttcgcgcaagta 1500  
Qy 1501 gggaaagtttgagttcttgagttatccacgcgcaatagtgtaaatgacttctgtaaacg 1560  
Db 1501 gggaaagtttgagttcttgagttatccacgcgcaatagtgtaaatgacttctgtaaacg 1560  
Qy 1561 atgtgcaatatagcgttaagactatgaataacacgcgtcgagacgcgtcgaagaacgcg 1620  
Db 1561 atgtgcaatatagcgttaagactatgaataacacgcgtcgagacgcgtcgaagaacgcg 1620  
Qy 1621 gtgtggcgacgcgcaacactacactgcgtcgctaaagaacggttaaatcttcggttaaaaa 1680  
Db 1621 gtgtggcgacgcgcaacactacactgcgtcgctaaagaacggttaaatcttcggttaaaaa 1680  
Qy 1681 atgaatctcgggcatgggttaataagctcctgcagaaatgcacagaaatgttccctccat 1740  
Db 1681 atgaatctcgggcatgggttaataagctcctgcagaaatgcacagaaatgttccctccat 1740  
Qy 1741 caaagaataacacgcgaacacactcaacgcgcaagttatgtaagcgtgataaacacatg 1800  
Db 1741 caaagaataacacgcgaacacactcaacgcgcaagttatgtaagcgtgataaacacatg 1800  
Qy 1801 aaatgacctcagaatactgcgcatatgagcgtggaagttccggttaacggtcttat 1860  
Db 1801 aaatgacctcagaatactgcgcatatgagcgtggaagttccggttaacggtcttat 1860  
Qy 1861 ctgagtcagagagatgcgcgaacaatgctgcgacatgcccagcgtcttcaattcat 1920  
Db 1861 ctgagtcagagagatgcgcgaacaatgctgcgacatgcccagcgtcttcaattcat 1920  
Qy 1921 caccgatgagagaggaagacgcgccctcctcaaaaacaaagatggtggaagatatctgat 1980  
Db 1921 caccgatgagagaggaagacgcgccctcctcaaaaacaaagatggtggaagatatctgat 1980  
Qy 1981 cctgggttcagagacgttgccttaactgcgggaa 2015  
Db 1981 cctgggttcagagacgttgccttaactgcgggaa 2015

RESULT 4  
AA167035  
ID AA167035 standard; DNA; 8509 BP.  
XX  
XX AA167035:  
AC  
XX  
XX  
DT 11-FEB-2002 (first entry)  
DE  
XX  
XX Nucleotide sequence of an endogenous ketoglulonigenium plasmid.  
KW Ketoglulonigenium; plasmid; strain NRRL B-30035; vector; ds.  
XX  
XX OS Synthetic.  
XX  
XX W0200177159-A2.  
XX  
XX  
XX 18-OCT-2001.  
XX  
XX



XX	05-APR-2001; 2001WO-US11058.
PF	
PR	05-APR-2000; 2000US-194624P.
XX	
PA	(UNMS ) UNIV MICHIGAN STATE
PA	(ARCH ) ARCHER-DANIELS MIDLAND CO.
XX	
FI	Schmidt TM, Stoddard SF;
DR	WPI; 2001-657165/75.
XX	
PT	New nucleic acid comprising the sequence of a Ketogulonigenium plasmid
PT	designated pADM291 is endogenous to microorganism strain NRRL B-30035
XX	
XX	
PS	Claim 1; Fig 1A-E; 14pp; English.
XX	
CC	The present sequence represents a purified nucleic acid of a
CC	Ketogulonigenium plasmid endogenous to strain NRRL B-30035. The
CC	polynucleotide can be used in a vector. The vector comprises either
CC	(i) the polynucleotide and a marker gene or (ii) the polynucleotide, a
CC	promoter, a transcription terminator and a discrete series of restriction
CC	endonuclease recognition sites between the promoter and transcription
XX	terminator.
XX	
Sequence	8509 BP; 2065 A; 2239 C; 2345 G; 1860 T; 0 other;

[illegible]

Db	3555	cgcctccagcgctccaaagctccatgacatttaatgataagccacctcgcgcgccagctgctc	3614
Oy	661	gacgacgtgbcgcccctgaaatgacgctgacgcgaacatctgcgaatctgaacgcgatgaaac	720
Db	3615	gacgacgtgbcgcccctgaaatgacgctgacgcgaacatctgcgaatctgaacgcgatgaaac	3674
Oy	721	catgacacgtgaaagacatgcaccccgctgctcgagagagctaaagctcgcgctgttgaacct	780
Db	3675	catgacacgtgaaagacatgcaccccgctgctcgagagagctaaagctcgcgctgttgaacct	3734
Oy	781	gacgacccctgcgaataatgacatgcgtgacacgtcgcgcgctctgtgacatgaaagcgcgaatagac	840
Db	3735	gacgacccctgcgaataatgacatgcgtgacacgtcgcgcgctctgtgacatgaaagcgcgaatagac	3794
Oy	841	taccgcgaagacgaacagcgcgacgaacccctacgtgaaacgtgaaacccctcccgagatccatcgt	900
Db	3795	taccgcgaagacgaacagcgcgacgaacccctacgtgaaacgtgaaacccctcccgagatccatcgt	3854
Oy	901	cgtatgctcgcgagatctgcgaacacatctgcacatcttcgaacgcgtcaaacgctatccatcgt	960
Db	3855	cgtatgctcgcgagatctgcgaacacatctgcacatcttcgaacgcgtcaaacgctatccatcgt	3914
Oy	961	ggtacgaatgattccgctgctgctgcttcgaacgcgtctctagctctgcgaatcttgatcgtg	1020
Db	3915	ggtacgaatgattccgctgctgctgcttcgaacgcgtctctagctctgcgaatcttgatcgtg	3974
Oy	1021	atbtagcgcgaacaaaccttlaacgctcccccagcttgcgcgcgcgtcctctgaaatgcccgaagga	1080
Db	3975	atbtagcgcgaacaaaccttlaacgctcccccagcttgcgcgcgcgtcctctgaaatgcccgaagga	4034
Oy	1081	aagatgcttcgcttgcgaacgaagcttlaacgaattgtcttccaacccctgacatctgatatgctc	1140
Db	4035	aagatgcttcgcttgcgaacgaagcttlaacgaattgtcttccaacccctgacatctgatatgctc	4094
Oy	1141	aacacattatccgctctgcgaacatctgaacgcgaacgcgcgaacaaagattggccgtatgctgga	1200
Db	4095	aacacattatccgctctgcgaacatctgaacgcgaacgcgcgaacaaagattggccgtatgctgga	4154
Oy	1201	agctgcgacatacagctctggaagctgaaagacgcgaacccacgcgtcgccagcgctgagctgcg	1260
Db	4155	agctgcgacatacagctctggaagctgaaagacgcgaacccacgcgtcgccagcgctgagctgcg	4214
Oy	1261	ggttccaaagctcgctgcgcgaagatgcctctgcgaagagggcgacggaacacgaataagccccctcc	1320
Db	4215	ggttccaaagctcgctgcgcgaagatgcctctgcgaagagggcgacggaacacgaataagccccctcc	4274
Oy	1321	tccccgaagcgcgcgcgcgaatctaacctacgaatcccaagcttctgcgtgcgaagctgcgaacgcctctc	1380
Db	4275	tccccgaagcgcgcgcgcgaatctaacctacgaatcccaagcttctgcgtgcgaagctgcgaacgcctctc	4334
Oy	1381	ggcgcgaacaaagacaaacgaatctgcgtgcgaacatctgcgcgcgcgcgtcttctgcgtgcgaagga	1440
Db	4335	ggcgcgaacaaagacaaacgaatctgcgtgcgaacatctgcgcgcgcgcgtcttctgcgtgcgaagga	4394
Oy	1441	ggcgcgcgcctcgcgaacgcgtgcgaacacatccgaacaaacgcgtttttatgattctcgcgcgaagga	1500
Db	4395	ggcgcgcgcctcgcgaacgcgtgcgaacacatccgaacaaacgcgtttttatgattctcgcgcgaagga	4454
Oy	1501	ggcgaagcgttcgagcttctgaagatattccaccgcgaatagtgtaaatgacattctgcgaacg	1560
Db	4455	ggcgaagcgttcgagcttctgaagatattccaccgcgaatagtgtaaatgacattctgcgaacg	4514
Oy	1561	atgcgcacatactgcgtgaagactatgaaatacaacgcgcgcgcgaacgcgcgcgaacaaagcgaag	1620
Db	4515	atgcgcacatactgcgtgaagactatgaaatacaacgcgcgcgcgaacgcgcgcgaacaaagcgaag	4574
Oy	1621	gtctgtgcgcacgcgaacacatccatctgcgcgcgttaaaagcgtgaaatcttcgcgttaaaaag	1680
Db	4575	gtctgtgcgcacgcgaacacatccatctgcgcgcgttaaaagcgtgaaatcttcgcgttaaaaag	4634
Oy	1681	atgaatctcgggcgcgtggtctatagactctgcgcgaatctgcacagcgttctctcccatct	1740



Db	4635	atgaatctggggcatlvggtattatgattccctgcgaatctgcacagagtggttccctccatt	4654
Oy	1741	caagaatacacccgaacacactaacacgacgacatlatatgtaagtcgtatanaacacatg	1800
Db	4695	caagaataatcacccgaacacactaacacgacgacatlatatgtaagtcgtatanaacacatg	4754
Oy	1801	aaatgacctagaatatcagacatcttagagcgttgaattcggactttagcgatctttat	1860
Db	4755	aaatgacctagaatatcagacatcttagagcgttgaattcggactttagcgatctttat	4814
Oy	1861	ctgatgccaaaggagatcgcgacaacatlgcgcgacatlgccgaagcgtcttcaattcat	1920
Db	4815	ctgatgccaaaggagatcgcgacaacatlgcgcgacatlgccgaagcgtcttcaattcat	4874
Oy	1921	caacgatgagagaggaagacgcgcgcctcccaaaaacaaagatcgttggagaatattcgaat	1980
Db	4875	caacgatgagagaggaagacgcgcgcctcccaaaaacaaagatcgttggagaatattcgaat	4934
Oy	1981	ctctggcttcaagagccttgccttactatcgcgaa	2015
Db	4935	ctctggcttcaagagccttgccttactaaacccgaa	4969

RESULT	5
AA518307	
ID	AA518307 standard; DNA; 8509 BP.

AC AAS18307

DT 12-MAR-2002 (first entry)

DE DNA sequence of *Ketogulonigenium* endogenous plasmid PADM291.

KW Cloning vector; Ketogulonigenium replicon; endogenous plasmid;

KW transformed host cell; Escherichia coli; pADM291; circular; cyclic; ds.

OS Ketogulonigenium sp. strain ADM291-19.

PN WO200177347-A2.

PD 18-OCT-2001.

PF 05-APR-2001; 2001WO-US11059.

PR 05-APR-2000; 2000US-194625P.

PA (ARCH ) ARCHER-DANIELS MIDLAND CO.

PA (DELI/) D'ELIA J.

PI D'Elia J;

DR WPI; 2002-049150/06

PT Novel nucleic acid vector comprising Ketogulonigenium replicon found on

PT polypeptides and/or transcripts by culturing host cells transformed

XX 11

PS Example 7; Fig 2; 66pp; English.

CC The present invention relates to the isolation of vectors comprising

CC The invention also describes methods of transforming host cells with

culturing the transformed host cells. The vectors are useful for

CC The vectors which have a replicon functional in both Ketopw  
CC transforming a host cell by conjugation or electroporation  
CC

CC and *Escherichia coli*, enable the cloning of certain genes of  
CC Ketoulonigenium in *E. coli* as the latter is an efficient host for

cc Ketocyclonigenium endogenous plasmid PADM291.

Sequence 8509 BP: 2065 A: 2239 C: 2345 G: 1860 T: 0 other:

2

Query Match	95.0%;	Score 2007;	DB 24;	Length 8509;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2010; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

OY	1	ggatgtagtgcgnaattcatagaagatttgtagtaggtgtagtgtagcgccctctgaacggggtg	60
Db	2955	ggcaaatgggttcgnaattcatagaattcttgtagagtgtagcgccctctgaacggggtg	3014
OY	61	ctgcgcggaatcctctgtcttcacgtacagggcgacaattgagaggtgttattgtccccctg	120
Db	3015	ctgcgcggaatcctctgtcttcacgtacagggcgacaattgagaggtgttattgtccccctg	3074
OY	121	tatgcctctctgcgttgtagcatcttggttcatctctgcgccggactatgatacttcgttag	180
Db	3075	tatgcctctctgcgttgtagcatcttggttcatctctgcgccggactatgatacttcgttag	3134
OY	181	gattacatgatagtcttcgcctctgcggcgtgtcttcgcgcctgttcgcggcgtgttcgc	240
Db	3135	gattacatgatagtcttcgcctctgcggcgtgtcttcgcgcctgttcgcggcgtgttcgc	3194
OY	241	gggcctgtccctctgttcgcgcgcgttcctcaacttttcaacaataatgtagcgaagc	300
Db	3195	gggcctgtccctctgttcgcgcgcgttcctcaacttttcaacaataatgtagcgaagc	3254
OY	301	cctctctgtctatagttctctatagttctacagacaattacacataatcatatagatt	360
Db	3255	cctctctgtctatagttctctatagttctacagacaattacacataatcatatagatt	3314
OY	361	attcgcttaaaaggagtaattgtagccgcgaaaaggagtaattgtagccgcgaaaaggagt	420
Db	3315	attcgcttaaaaggagtaattgtagccgcgaaaaggagtaattgtagccgcgaaaaggagt	3374
OY	421	aattgcggccgaaaaggagtaattgtagccgatactggttcttcaatggggaggaatcc	480
Db	3375	aattgcggccgaaaaggagtaattgtagccgatactggttcttcaatggggaggaatcc	3434
OY	481	ccttaatcatcttctcccatgtggaaagacaacaagtgtagccgcgaacggccttcgac	540
Db	3435	ccttaatcatcttctcccatgtggaaagacaacaagtgtagccgcgaacggccttcgac	3494
OY	541	cagacaaaactgtgtgtccctctgcgcggagtgtgcagagaggggtctatatgcgcaatccgc	600
Db	3455	cagacaaaactgtgtgtccctctgcgcggagtgtgcagagaggggtctatatgcgcaatccgc	3554
OY	601	cgccctgcagcgctcacaagctcatatgcatlttaatgatatgacactgtgcggcgcgcatagtc	660
Db	3555	cgccctgcagcgctcacaagctcatatgcatlttaatgatatgacactgtgcggcgcgcatagtc	3614
OY	661	gatatgtagtgcgcataatgaatgtgcgtctgtgcgcgaatacttcgcgcgaatcgcgaatgaanaac	720
Db	3615	gatatgtagtgcgcataatgaatgtgcgtctgtgcgcgaatacttcgcgcgaatcgcgaatgaanaac	3674
OY	721	catgaacgttgtagaactcgcgaactcgcgtctgttcagaaagactagccgcgtgtgcgttgttgaacat	780
Db	3675	catgaacgttgtagaactcgcgaactcgcgtctgttcagaaagactagccgcgtgtgcgttgttgaacat	3734
OY	781	gattgaaccttgaaaagatgatacgttggaacagtgtagcgttgttgtagtagagcggaataagac	840
Db	3735	gattgaaccttgaaaagatgatacgttggaacagtgtagcgttgttgtagtagagcggaataagac	3794
OY	841	taccgcgcagagggagcaagcgcggaactctctagtgtagcgttgtagccttccggagataatccgt	900
Db	3795	taccgcgcagagggagcaagcgcggaactctctagtgtagcgttgtagccttccggagataatccgt	3854
OY	901	cgtagtagggggaggttgtagaaccaacttgtagccattctgacacgcgcgaaggtattcatatc	960
Db	3855	cgtagtagggggaggttgtagaaccaacttgtagccattctgacacgcgcgaaggtattcatatc	3914
OY	961	ggtagtaagtaattccgcgtgcctgttcttcacgacagctctcagttctgcgcaactcttgatcgg	1020
Db	3915	ggtagtaagtaattccgcgtgcctgttcttcacgacagctctcagttctgcgcaactcttgatcgg	3974



```

OY 1021 atgagcgcgaaaaaccttaacggtcccccaggttgccggtcctcttggaatgcccagagga 1080
    |||||
DB 3975 atgagcgcgaaaaaccttaacggtcccccaggttgccggtcctcttggaatgcccagagga 4034
OY 1081 aagaatgcttcgttggaacgacgttaacacagatttgcctcaaacctgcactgtgatgagtc 1140
    |||||
DB 4035 aagaatgcttcgttggaacgacgttaacacagatttgcctcaaacctgcactgtgatgagtc 4094
OY 1141 aaccattatcgcgtctgacatttgaacgcgaacccgaacgaatttgccgttaacgttcgga 1200
    |||||
DB 4095 aaccattatcgcgtctgacatttgaacgcgaacccgaacgaatttgccgttaacgttcgga 4154
OY 1201 aagtgtaactaagctctggaagtgaaagacgacccaacgctgcgcagcgcgagctggcg 1260
    |||||
DB 4155 aagtgtaactaagctctggaagtgaaagacgacccaacgctgcgcagcgcgagctggcg 4214
OY 1261 ggttcacaaggtcgtgtagatgctcgtcgcagaagggcgagcggaagaacgaatagccctcc 1320
    |||||
DB 4215 ggttcacaaggtcgtgtagatgctcgtcgcagaagggcgagcggaagaacgaatagccctcc 4274
OY 1321 ttcccaagaagcgcgagatcaccatcaccatcgttgctgagctggaacgctcgtc 1380
    |||||
DB 4275 ttcccaagaagcgcgagatcaccatcaccatcgttgctgagctggaacgctcgtc 4334
OY 1381 ggcagcaacaaggaacacgacatcgtcgtcagacttcgcggcgttctcgtcgggaaga 1440
    |||||
DB 4335 ggcagcaacaaggaacacgacatcgtcgtcagacttcgcggcgttctcgtcgggaaga 4394
OY 1441 ggcgtgctcgtcggcgctgcgaacacatcgaataaactgtttttagatttctgcgcaaga 1500
    |||||
DB 4395 ggcgtgctcgtcggcgctgcgaacacatcgaataaactgtttttagatttctgcgcaaga 4454
OY 1501 gggaaagtttgagtttcttgagttatttcacgcgaatagtgtaaatgacttcgtggaacg 1560
    |||||
DB 4455 gggaaagtttgagtttcttgagttatttcacgcgaatagtgtaaatgacttcgtggaacg 4514
OY 1561 atgtgcataatagcgtgtaagacttgaataacacgcgtcgcaacgctgcaaaaacgcaacg 1620
    |||||
DB 4515 atgtgcataatagcgtgtaagacttgaataacacgcgtcgcaacgctgcaaaaacgcaacg 4574
OY 1621 ggtgtgcgacccgcaacacatcactcgtcggtctlaaaaagcgtlaaaaattcccggtlaaaaag 1680
    |||||
DB 4575 ggtgtgcgacccgcaacacatcactcgtcggtctlaaaaagcgtlaaaaattcccggtlaaaaag 4634
OY 1681 atgaatctggggcagatgggttataagatcctcgtcagaatttgacagagatgttctcccat 1740
    |||||
DB 4635 atgaatctggggcagatgggttataagatcctcgtcagaatttgacagagagatgttctcccat 4694
OY 1741 caaaagaatataccgcaaacacactaatacgcgaagtalatgttaagcgtgaatgaacaacatg 1800
    |||||
DB 4695 caaaagaatataccgcaaacacactaatacgcgaagtalatgttaagcgtgaatgaacaacatg 4754
OY 1801 aaatgacctcaagaatacgcgcatcagcgtggaagtgttgagctttagcgcgttctat 1860
    |||||
DB 4755 aaatgacctcaagaatacgcgcatcagcgtggaagtgttgagctttagcgcgttctat 4814
OY 1861 ctgatacgaaggagatcgcgacaatgycgcacatgycgcacatgycgcgcgtcttcaattcat 1920
    |||||
DB 4815 ctgatacgaaggagatcgcgacaatgycgcacatgycgcacatgycgcgcgtcttcaattcat 4874
OY 1921 caccgatgagagaggaagacgcgcgcgcctcaaaaaaagaatgtgtgaagatattctgat 1980
    |||||
DB 4875 caccgatgagagaggaagacgcgcgcgcctcaaaaaaagaatgtgtgaagatattctgat 4934
OY 1981 cctggggttcagagacgttgccttactgtgcgga 2015
    |||||
DB 4935 cctggggttcagagacgttgccttactgtgcgga 4969

```

RESULT 6  
AAQ21036  
ID AAQ21036 standard; DNA: 731 BP.

```

XX XX AAQ21036:
AC XX 29-MAY-1992 (first entry)
DT XX
XX XX
DE XX Region specific to Mycoplasma hyopneumoniae.
XX XX ribosomal RNA: detection kit; polymerase chain reaction; PCR; ss.
XX XX
OS Mycoplasma hyopneumoniae.
XX XX
XX XX JP04004899-A.
XX XX
XX XX 09-JAN-1992.
XX XX
XX XX 24-APR-1990; 90JP-0106354.
XX XX
XX XX 24-APR-1990; 90JP-0106354.
XX XX
XX XX (TAKA-) TAKARA SHUZO KK.
XX XX
XX XX WPI; 1992-060511/08.
XX XX
XX XX Detecting mycoplasma - by detecting and amplifying DNA sequence
PT coding for mycoplasma and RNA
PT
XX XX Example 2; Fig 9; 13pp; Japanese.
XX XX
XX XX This region of M.hyopneumoniae-specific DNA was amplified by PCR.
CC See also AAQ21028-Q21051.
XX XX
XX XX Sequence 731 BP; 252 A; 94 C; 103 G; 259 T; 23 other;

```

Query Match 1.9%; Score 40.4; DB 13; Length 731;  
Best Local Similarity 49.5%; Pred. No. 0.081;  
Matches 104; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```

OY 1593 acggctcgacagcgtcgaacgaacgagtggtggcgacccgcaacatcactcgtggcgcta 1652
    |||||
DB 421 agggctctaaagcctcctaacaacaaatgcgaagaagcaaacagctcttcaaaa 480
OY 1653 aaagcgtaaatcttcggttaaaaagatgtaattcgtggcagatggttatagatcgtga 1712
    |||||
DB 481 ctgaataatcaatttatctcgtatgataaatacaggtaatatctcctaagtatcttaa 540
OY 1713 gaatgcacagagtggttctcccatcctaagaataacacggaacacactaacaacgcaa 1772
    |||||
DB 541 aaatnaaccgagttatcttgccttttaatttttgaaataaaaaataaagaaga 600
OY 1773 gtatatgtaagcgtgatgaacaacatgaa 1802
    |||||
DB 601 gattatgcataaatgttaaaataccttaa 630

```

RESULT 7  
AAS52512  
ID AAS52512 standard; DNA: 1980 BP.  
AC AAS52512;  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX  
XX E. coli DNA for cellular proliferation protein #234.  
XX  
XX Antisense; ds; prokaryotic cellular proliferation gene;  
XX antibiotic; antibacterial; drug design.  
XX Escherichia coli.  
XX  
XX WO200170955-A2.  
XX  
XX 27-SEP-2001.  
PD















DT	12-MAR-2002	(first entry)	
XX			
XX		Replicon sequence #2 of Ketogulonigenium endogenous plasmid pADMX6L1.	
DE			
XX		Ketogulonigenium; 2-keto-L-gulononic acid; L-sorbose; sorbitol;	
KW		plasmid replicon pADMX6L1; replicon; ds.	
XX			
OS		Ketogulonigenium robustum strain pADMX6L.	
XX			
PN		WO200177348-A2.	
XX			
PD		18-OCT-2001.	
XX			
XX			
PF	05-APR-2001;	2001WO-US11097.	
XX			
PR	05-APR-2000;	2000US-194627P.	
XX			
PA	(ARCH )	ARCHER-DANIELS MIDLAND CO.	
PA	(DELI/ )	D'ELIA J.	
PA	(STOD/ )	STODDARD S F.	
XX			
PI	D'Elia J,	Stoddard SF;	
XX			
DR	WPI: 2002-041295/05.		
XX			
XX			
PT	New bacterium of Ketogulonigenium genus, useful for producing		
PT	2-keto-L-gulononic acid from sorbose or sorbitol, comprises transgene		
PT	containing DNA sequence from endogenous Ketogulonigenium plasmid -		
XX			
PS	Claim 15; Page 97; 116pp; English.		
XX			
CC	The present invention relates to a new bacterium of genus		
CC	Ketogulonigenium. Ketogulonigenium may further comprise a		
CC	transgene, comprising a DNA sequence from an endogenous		
CC	Ketogulonigenium plasmid. Methods for transforming		
CC	Ketogulonigenium are also described. The invention is useful for		
CC	producing 2-keto-L-gulononic acid (2-KLG) from L-sorbose or sorbitol.		
CC	The present sequence represents the nucleotide sequence of		
CC	replicon #2 on the Ketogulonigenium endogenous plasmid pADMX6L1.		
CC	Note: The present sequence for SEQ ID No 5 given in the sequence		
CC	listing is different from that given for SEQ ID No 5 in		
CC	Fig 5 (AAS17123).		
SO			
	Sequence 1458 BP; 365 A; 423 C; 415 G; 254 T; 1 other:		
	Query Match	1.8%;	Score 37.4; DB 24; Length 1458;
	Best Local Similarity	54.8%;	Pred. No. 1.1;
	Matches 74; Conservative	0;	Mismatches 61; Indels 0; Gaps 0;
Qy	1603 aggtgcacaagaacgaaggtgtggtgcacgcacacatactactcgtggcgctcaaaaagcgta	1662	
Db	986 agcggtlaaacacacttgatgtgttccggcgcaacctgcataaaccttaacttgaca	1045	
Qy	1663 aaattccggtlaaaaagaatgaatcctggggcatcgggttataagatcctgagaatigcaca	1722	
Db	1046 agattccagtggttcgagtgatgcagaagacgltgacacatatagaccctcagatgacaa	1105	
Qy	1723 gaagtgttcctccca	1737	
Db	1106 gagttaccagccaa	1120	
	RESULT 14		
	AAS17119		
ID	AAS17119 standard; DNA; 7029 BP.		
XX			
AC	AAS17119;		
XX			
DT	12-MAR-2002 (first entry)		
XX			
DE	Ketogulonigenium endogenous plasmid pADMX6L1 DNA sequence.		
XX			



KM ketogulonigenium; 2-keto-L-gulononic acid; L-sorbose; sorbitol;  
KM plasmid PADMX6L1; replication protein; cyclic; circular; ds.  
XX  
OS Ketogulonigenium robustum strain ADMX6L.  
XX  
XX Key Location/Qualifiers  
FH CDS 2255..2710  
FT /tag- a  
FT /product- "PADMX6L1 replication protein"  
FT /note- "PADMX6L1 rep ORF"  
FT 1755..3210  
FT misc-feature  
FT /tag- c  
FT /note- "PADMX6L1 replicon, corresponds to AAS17123"  
XX  
XX WO200177348-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 05-APR-2001; 2001WO-US11097.  
XX  
XX 05-APR-2000; 2000US-194627P.  
XX  
XX (ARCH ) ARCHER-DANIELS MIDLAND CO.  
XX (DELI/) D-ELIA J.  
XX (STOD/) STODDARD S F.  
XX  
XX D'Elia J, Stoddard SF.  
XX  
XX WPI: 2002-041295/05.  
XX P-PSDB: AAU11038.  
XX  
XX New bacterium of Ketogulonigenium genus, useful for producing  
PT 2-keto-L-gulononic acid from sorbose or sorbitol, comprises transgene  
PT containing DNA sequence from endogenous Ketogulonigenium plasmid -  
XX  
XX Claim 10; Fig 1; 116pp; English.  
XX  
XX The present invention relates to a new bacterium of genus  
CC Ketogulonigenium. Ketogulonigenium may further comprise a  
CC transgene, comprising a DNA sequence from an endogenous  
CC Ketogulonigenium plasmid. Methods for transforming  
CC Ketogulonigenium are also described. The invention is useful for  
CC producing 2-keto-L-gulononic acid (2-KLG) from L-sorbose or sorbitol.  
CC The present sequence represents Ketogulonigenium endogenous plasmid  
CC PADMX6L1 DNA.  
XX  
XX Sequence 7029 BP; 1455 A; 2236 C; 2094 G; 1184 T; 0 other;  
SQ  
Query Match 1.8%; Score 37.4; DB 24; Length 7029;  
Best Local Similarity 54.8%; Pred. No. 2.7;  
Matches 74; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

XX  
XX Pseudomonas aeruginosa DNA for cellular proliferation protein #309.  
DE  
XX  
XX Antisense; ds; prokaryotic cellular proliferation gene;  
KM antibiotic; antibacterial; drug design.  
XX  
XX Pseudomonas aeruginosa.  
OS  
XX  
XX WO200170955-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 21-MAR-2001; 2001WO-US09180.  
XX  
XX  
XX 21-MAR-2000; 2000US-191078P.  
PR 23-MAY-2000; 2000US-205848P.  
PR 26-MAY-2000; 2000US-207727P.  
PR 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
XX Yamamoto RT, Xu HH;  
XX  
XX WPI: 2001-611495/70.  
XX P-PSDB: AAU36319.  
XX  
XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -  
XX  
XX Claim 27; Seq ID No 7815; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence encodes an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
XX Sequence 858 BP; 149 A; 294 C; 291 G; 124 T; 0 other;  
SQ  
Query Match 1.7%; Score 36.8; DB 23; Length 858;  
Best Local Similarity 43.6%; Pred. No. 1.2;  
Matches 164; Conservative 0; Mismatches 212; Indels 0; Gaps 0;







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 15:21:38 ; Search time 59.35 Seconds  
(Without alignments)  
8740.993 Million cell updates/sec

Title: US-09-826-206-1  
Perfect score: 2112  
Sequence: 1 ggcattgggtcgaatcatc.....tgaggaaatagggaacgtg 2112

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/2/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/lna/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
C 1	45.8	2.2	7218	1	US-08-232-463-14		Sequence 14, App1
2	37.4	1.8	7218	1	US-08-232-463-14		Sequence 14, App1
3	36	1.7	289	4	US-09-007-005-17		Sequence 17, App1
4	36	1.7	289	4	US-09-244-796-17		Sequence 17, App1
C 5	35.8	1.7	4403765	4	US-09-103-840A-2		Sequence 2, App1
C 6	35.8	1.7	4411529	4	US-09-103-840A-1		Sequence 1, App1
7	34.4	1.6	500	3	US-09-141-000-2		Sequence 2, App1
C 8	34.4	1.6	5176	4	US-09-182-024A-1		Sequence 1, App1
C 9	34.2	1.6	1643	2	US-08-486-839-3		Sequence 3, App1
C 10	34.2	1.6	1643	3	US-09-151-011-3		Sequence 3, App1
C 11	34.2	1.6	1643	4	US-09-343-623-3		Sequence 3, App1
C 12	34.2	1.6	1713	2	US-08-486-839-5		Sequence 5, App1
C 13	34.2	1.6	1713	3	US-09-151-011-5		Sequence 5, App1
C 14	34.2	1.6	1713	4	US-09-343-623-5		Sequence 5, App1
C 15	33.6	1.6	2756	1	US-08-187-793-1		Sequence 1, App1
C 16	33.6	1.6	3318	1	US-08-187-793-3		Sequence 3, App1
C 17	33.6	1.6	4104	1	US-07-998-003A-94		Sequence 94, App1
C 18	33.6	1.6	4104	1	US-08-453-274B-94		Sequence 94, App1
19	33.6	1.6	4104	1	US-08-453-695A-94		Sequence 94, App1
20	33.6	1.6	4104	1	US-08-268-161A-94		Sequence 94, App1
21	33.6	1.6	4104	2	US-08-453-702A-94		Sequence 94, App1
22	33.6	1.6	4104	4	US-09-099-639-94		Sequence 94, App1
23	33.6	1.6	4104	5	PCT-US93-12588-94		Sequence 94, App1
24	33.6	1.6	4104	5	PCT-US95-08071-94		Sequence 94, App1
25	33.6	1.6	4650	1	US-07-998-003A-102		Sequence 102, App
26	33.6	1.6	4650	1	US-08-453-274B-102		Sequence 102, App
27	33.6	1.6	4650	1	US-08-453-695A-102		Sequence 102, App

28	33.6	1.6	4650	1	US-08-268-161A-102		Sequence 102, App
29	33.6	1.6	4650	2	US-08-453-702A-102		Sequence 102, App
30	33.6	1.6	4650	4	US-09-099-639-102		Sequence 102, App
31	33.6	1.6	4650	5	PCT-US93-12588-102		Sequence 102, App
32	33.6	1.6	4650	5	PCT-US95-08071-102		Sequence 102, App
33	33.6	1.6	2473	4	US-09-173-914-3		Sequence 3, App1
34	33.2	1.6	3974	4	US-08-467-504-3		Sequence 3, App1
35	33	1.6	3200	1	US-08-444-405-1		Sequence 1, App1
36	33	1.6	3200	1	US-08-384-850-1		Sequence 1, App1
C 37	32.6	1.5	5181	1	US-08-257-073-10		Sequence 10, App1
C 38	32.4	1.5	68750	3	US-09-335-409-1		Sequence 1, App1
C 39	32.4	1.5	68750	4	US-09-568-102-1		Sequence 1, App1
C 40	32.4	1.5	68750	4	US-09-567-969-1		Sequence 1, App1
C 41	32.4	1.5	68750	4	US-09-568-480-1		Sequence 1, App1
C 42	32.4	1.5	68750	4	US-09-568-486-1		Sequence 1, App1
C 43	32.4	1.5	68750	4	US-09-568-472-1		Sequence 1, App1
C 44	32.4	1.5	71989	4	US-09-443-501A-2		Sequence 2, App1
C 45	32.2	1.5	1636	4	US-09-039-198A-1		Sequence 1, App1

## ALIGNMENTS

RESULT 1  
US-08-232-463-14/C  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
ATTORNEY/AGENT INFORMATION:  
FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-Fls  
US-08-232-463-14



Query Match	Score	DB	Length
Best Local Similarity	3.8%	Pred. No. 0.000221	
Matches	14	Conservative	205
		Mismatches	152
		Indels	0
		Gaps	0

US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEFFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pT9pt-F15
US-08-232-463-14

```

Query Match	1.88;	Score 37.4;	DB 1;	Length 7218;
Best Local Similarity	9.38;	Pred. No. 0.15;		
Matches	26;	Conservative 136;	Mismatches 117;	Indels 0;
				Gaps 0

Qy	111	ttggcccttgatcgtctctcgttgagcgaattggatcactcgtccggagcaatgat	170
Db	1213	yy	1272
Qy	171	tcgcctagaggaatactacgatattctcgcctgcggcctgtcggagctgtcggcgtgtg	230
Db	1273	yy	1332
Qy	231	cgagcctgtcgggcctgcctccctctgtcccgcgctgcctccacttttcacatcaaaaa	290
Db	1333	yy	1392
Qy	291	tggcggaagccctctctgttctataagttcttatagttcataagaataataccacataat	350
Db	1393	yytctaccaaatcttctatc	1452
Qy	351	tcgaatagcttatcgtctaaagaagaaattggcgccg	389
Db	1453	tcttttactacttgcatgataagataggaaattacagtgatgc	1491

```

US-09-007-005-17
? RESULT 3
? Sequence 17, Application US/09007005B
? Patent No. 6258558
? GENERAL INFORMATION:
? APPLICANT: Szostak, Jack W.
? APPLICANT: Roberts, Richard W.
? APPLICANT: Liu, Rihе
? TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
? TITLE OF INVENTION: FUSIONS
? FILE REFERENCE: 00786/350003
? CURRENT APPLICATION NUMBER: US/09/007, 005B
? CURRENT FILING DATE: 1998-01-14
? EARLIER APPLICATION NUMBER: 60/035, 963
? EARLIER FILING DATE: 1997-01-27
? EARLIER APPLICATION NUMBER: 60/064, 491
? EARLIER FILING DATE: 1997-11-06
? NUMBER OF SEQ ID NOS: 33
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 17
? LENGTH: 289
? TYPE: RNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Translation template
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(289)
? OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```

Query Match	1.7%	Score 36	DB 4	Length 289	
Best Local Similarity	6.4%	Pred. No. 0.042			
Matches	14	Conservative	113	Indels	0
				Gaps	0
OY	1152	gcgtctacatcgaagcaagccacaaagtctgcgtgcgtgcgaagtgtgacat	1211		
	1152	gcgtctacatcgaagcaagccacaaagtctgcgtgcgtgcgaagtgtgacat	1211		



Db 21 rcturatururarcrararurarcrarurgrnmsrnmrsnmrsnmrs 80  
Qy 1212 aggttggaagtgaagaccacccagctgcagagcgagctgggggttccaagt 1271  
Db 81 rnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrs 140  
Qy 1272 cggctgagatcgctgcagagggcgaggaacgaatagccccctccccaagc 1331  
Db 141 rnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrs 200  
Qy 1332 gggcggaatacactacagtcagctggctggagctga 1371  
Db 201 rnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrs 240

RESULT 4  
US-09-244-796-17  
; Sequence 17, Application US/09244796  
; Patent No. 6281344  
; GENERAL INFORMATION:  
; APPLICANT: Roberts, Jack W.  
; APPLICANT: Roberts, Richard W.  
; APPLICANT: Liu, Rih  
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
; FILE REFERENCE: 00786/350007  
; CURRENT APPLICATION NUMBER: US/09/244,796  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/035,963  
; EARLIER FILING DATE: 1997-01-27  
; EARLIER APPLICATION NUMBER: 60/064,491  
; EARLIER FILING DATE: 1997-11-06  
; EARLIER APPLICATION NUMBER: 09/007,005  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 289  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Translation template  
; NAME/KEY: misc\_feature  
; LOCATION: (1)-(289)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-244-796-17

Query Match 1.7%; Score 36; DB 4; Length 289;  
Best Local Similarity 6.4%; Pred. No. 0.042; Mismatches 113; Indels 0; Gaps 0;  
Matches 14; Conservative 93;  
Qy 1152 ggcctgacattgacggaagccgacaaagtggccgtgagctggcaagtgtgactat 1211  
Db 21 rcturatururarcrararurarcrarurgrnmsrnmrsnmrsnmrs 80  
Qy 1212 aggttggaagtgaagaccacccagctgcagagcgagctgggggttccaagt 1271  
Db 81 rnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrs 140  
Qy 1272 cggctgagatcgctgcagagggcgaggaacgaatagccccctccccaagc 1331  
Db 141 rnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrs 200  
Qy 1332 gggcggaatacactacagtcagctggctggagctga 1371  
Db 201 rnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrsnmrs 240

RESULT 5  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A

; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
US-09-103-840A-2

Query Match 1.7%; Score 35.8; DB 4; Length 4403765;  
Best Local Similarity 48.7%; Pred. No. 37;  
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
Qy 633 gatagccactgcgggcgagccgcatgctgatatgtgycgaatgagctggccga 692  
Db 4203897 GATTGCGCGGCGCCGCCGAGCGCGCGATGTGTGTGCTGACGTGATGCTTGA 4203838  
Qy 693 cattgcgcacatgcagcgacgaagaacacatgacccgtgagagccctgacccgcgtttcga 752  
Db 4203837 CATGAGCGGCTTGACGTGCTGCACAGTTCGCGCAGATCCAGCCGCGGTTCT 4203778  
Qy 753 ggaagctagccgctgcggtgtgacccatgatgacccctgcaagaatgacgtgacgtcg 812  
Db 4203777 ACTGCTACGCGCCAGGACCGCGGTGGGAAGATCTGATGCTGCTGACCGGGGGCGA 4203718  
Qy 813 cggcttgatcgatgagcg 831  
Db 4203717 CGACTACGTACCAAGCCG 4203699

RESULT 6  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 1.7%; Score 35.8; DB 4; Length 4411529;  
Best Local Similarity 48.7%; Pred. No. 37;  
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;







COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,839  
FILING DATE: 07 - June - 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baron, Ronald J.  
REGISTRATION NUMBER: 29,281  
REFERENCE/DOCKET NUMBER: 294-26  
TELEPHONE: (516) 822-3550  
TELECOMMUNICATION INFORMATION:  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
US-08-486-839-3

Query Match 1.6%; Score 34.2; DB 2; Length 1643;  
Best Local Similarity 51.7%; Pred. No. 0.62;  
Matches 78; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 1146 ttatcgctgtgacatgacgcaagaagccgaccagattggccgtgacgtgcaagtgt 1205  
Db 1608 TATTTTCAAGTGAAGAGGCGCAGCCAGAGAACCCAGAAAAAGAGCAAGCCTGAGA 1549  
Qy 1206 gactataggctgggaagtgaagaagcaccacgtgcggcgagcgagcggtgtc 1265  
Db 1548 GCAGAAAGCCTGGATTAAGGAAGACACAGAAAGCCTGCAGAGACCGATTGGCGCC 1489  
Qy 1266 caagtcgctgcagatgctctgcgcagaagg 1296  
Db 1488 CAGGAAACCAGAGGCGCAGCTGTAGAGTG 1458

RESULT 10  
US-09-151-011-3/c  
Sequence 3, Application US/09151011  
Patent No. 6057142  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Human Chitinase, Its Recombinant  
TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in  
TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann & Baron, LLP  
STREET: 6900 Jericho Turnpike  
CITY: Syoset  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11791  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/151,011  
FILING DATE: 10 - September - 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Morris, Robert C.  
REGISTRATION NUMBER: 42,910  
REFERENCE/DOCKET NUMBER: 294-32 DIV  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
US-09-151-011-3

Query Match 1.6%; Score 34.2; DB 3; Length 1643;  
Best Local Similarity 51.7%; Pred. No. 0.62;  
Matches 78; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 1146 ttatcgctgtgacatgacgcaagaagccgaccagattggccgtgacgtgcaagtgt 1205  
Db 1608 TATTTTCAAGTGAAGAGGCGCAGCCAGAGAACCCAGAAAAAGAGCAAGCCTGAGA 1549  
Qy 1206 gactataggctgggaagtgaagaagcaccacgtgcggcgagcgagcggtgtc 1265  
Db 1548 GCAGAAAGCCTGGATTAAGGAAGACACAGAAAGCCTGCAGAGACCGATTGGCGCC 1489  
Qy 1266 caagtcgctgcagatgctctgcgcagaagg 1296  
Db 1488 CAGGAAACCAGAGGCGCAGCTGTAGAGTG 1458

RESULT 11  
US-09-343-623-3/c  
Sequence 3, Application US/09343623  
Patent No. 6303118  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A human chitinase, its recombinant  
TITLE OF INVENTION: production, its use for decomposing chitin, its use  
TITLE OF INVENTION: In therapy or prophylaxis against infection diseases.  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann & Baron  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: United States of America  
ZIP: 11758  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/343,623  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,839  
FILING DATE: 07-June-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baron, Ronald J.  
REGISTRATION NUMBER: 29,281  
REFERENCE/DOCKET NUMBER: 294-26  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1643 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA







```
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-09-343-623-5
```

```
Query Match 1.6%; Score 34.2; DB 4; Length 1713;
Best Local Similarity 51.7%; Pred. No. 0.64; Mismatches 73; Indels 0; Gaps 0;
Matches 78; Conservative 0;

QY 1146 ttatcgcgtctgacatgacgcaagccgaagattgcccgtacgtgcaagtgt 1205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1679 TATTTCAGTGAAGAGGGGAGCCGAGAGACCCAGAAAAAGAGGCAAGCTGAGA 1620

QY 1206 gactatgcttggaagtgaagaagacccaacgcgtgcgaagcgtgcgagttc 1265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1619 GCAGAAAGCCTGATAAAGGACACAGAAAGCCTGACAGAGCCAGATTGCGGCCCC 1560

QY 1266 caaggtcgcgcgaagatgctgcgcgaaggg 1296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1559 CAGGAAAACCCAGAGGACAGGCTGTAGAGTG 1529

RESULT 15
; US-08-187-793-1/c
; Sequence 1, Application US/08187793
; Patent No. 5614371
; GENERAL INFORMATION:
; APPLICANT: POSNER, JEROME B.
; APPLICANT: DARNELL, ROBERT B.
; APPLICANT: FURNEAUX, HENRY M.
; TITLE OF INVENTION: ANTIGEN RECOGNIZED BY ANTIBODIES ASSOCIATED
; TITLE OF INVENTION: WITH PARANEOPLASTIC OPSOCLONDS AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 4
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER & DUNHAM
; STREET: 30 ROCKEFELLER PLAZA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,793
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 691,559
; FILING DATE: 25-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39227
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: (212) 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1029
; OTHER INFORMATION:
; US-08-187-793-1
```

```
Query Match 1.6%; Score 33.6; DB 1; Length 2756;
Best Local Similarity 57.7%; Pred. No. 1.5; Mismatches 44; Indels 0; Gaps 0;
Matches 60; Conservative 0;

QY 264 tgcctcaacttttcacaataaataaalgaggcagccctctgtctatagttctat 323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1595 TTTCACAGTTTAACTTTTAAATATTTTACAGTGTATTATGCAACTATATGCTTTT 1536

QY 324 agtcatacgaataatcacataatcatatagcttatcgtc 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1535 TGATCATTTTAAATTTAAACTTATTTCAAAATATTGTTTCT 1492
```

Search completed: August 20, 2002, 17:59:00  
Job time: 9442 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 13:00:17 ; Search time 1661.44 Seconds  
(without alignments)  
1715.146 Million cell updates/sec

Title: US-09-826-206-1

Sequence: 1 ggcaatgggtcgaatcatc.....tgaggaaataggagacgtg 2112

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estc2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vitc:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	64.6	3.1	442	9	AM839280 CM1-LT006
2	46	2.2	206	10	BF962526 PM1-NN120
3	45.8	2.2	304	10	BF960900 MR3-NN021
4	44.6	2.1	230	10	C55473 C55473 YUJ1
5	44.4	2.1	269	10	BF947280 MR3-NN021
6	44.4	2.1	312	10	BF947371 MR3-NN021
7	44	2.1	392	9	AM497369 g559a02.y
8	44	2.1	487	10	BJ204205 BJ204205
9	44	2.1	501	10	BJ183869 BJ183869
10	44	2.1	546	10	BJ19016 BJ19016
11	44	2.1	548	10	BJ197576 BJ197576
12	43.2	2.0	217	10	BI033579 BI033579
13	43.2	2.0	844	12	AZ675008 EMTHS57TF
14	43	2.0	1005	12	CNS07CYP
15	42.8	2.0	228	10	BF957901 PM1-NN120
16	42.6	2.0	241	10	BF947370 MR3-NN021
17	41.8	2.0	274	10	BI036852 MR4-NT014

C 18	41.8	2.0	299	10	BF947669	BF947669 MR3-NN021
19	41.4	2.0	307	10	BF947656	BF947656 MR3-NN021
20	41	1.9	506	10	BF602600	BF602600 267870 MA
21	41	1.9	763	10	BF61628	BF61628 963025C01
22	40.8	1.9	411	10	BI041264	BI041264 MR4-NT014
C 23	40.8	1.9	1115	12	CNS06TWN	AL414741 17 end of
C 24	40.6	1.9	442	9	AV626426	AV626426 AV626426
C 25	40.2	1.9	147	10	BF964400	BF964400 PM1-NN120
26	40.2	1.9	214	10	BF957676	BF957676 PM1-NN120
27	40.2	1.9	234	10	BF958958	BF958958 PM1-NN120
28	40.2	1.9	241	10	BF957677	BF957677 PM1-NN120
29	40.2	1.9	767	10	AU004726	AU004726 AU004726
30	40.2	1.9	1066	12	CNS03K12	AL248195 Tetradon
31	40	1.9	335	10	D66082	D66082 CELK073E3R
32	39.8	1.9	163	10	BF947330	BF947330 MR3-NN021
C 33	39.8	1.9	348	9	AM707713	AM707713 832012A03
C 34	39.8	1.9	372	10	BE352244	BE352244 894057H08
35	39.8	1.9	1101	12	CNS00370	AL064465 Drosophila
36	39.6	1.9	914	12	CNS00C2P	AL059740 Drosophila
37	39.6	1.9	1067	12	CNS04200	AL314601 Tetradon
38	39.6	1.9	1101	12	CNS05N2N	AL344840 Tetradon
39	39.4	1.9	266	10	R03644	R03644 PK09f11.r1
40	39.4	1.9	344	10	D65444	D65444 CELK056F9R
C 41	39.4	1.9	578	10	BJ100728	BJ100728 BJ100728
C 42	39.4	1.9	619	9	AM152730	AM152730 JALSL3C10
C 43	39.2	1.9	353	10	BI033664	BI033664 PM1-NN120
C 44	39.2	1.9	534	12	BH332756	BH332756 CH230-201
45	39.2	1.9	601	12	BH282202	BH282202 CH230-205

#### ALIGNMENTS

RESULT 1	442 bp	MRNA	linear	EST 18-MAY-2000
AM839280/C				
LOCUS	CM1-LT0067-280100-109-h11	LT0067	Homo sapiens cDNA	MRNA sequence.
DEFINITION	AM839280			
ACCESSION	AM839280.1	GI:7933254		
VERSION				
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 442) Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W., Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.			
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)			
MEDLINE	20202663			
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?ftl=6t2-CM1-LT0067-280100-109-h11&ts=2000-01-28&ta=1) Seq primer: puc 18 forward High quality sequence start: 37 High quality sequence stop: 442. Location/Qualifiers 1..442 /organism="Homo sapiens"			



```

/db_xref="taxon:9606"
/clone_lib="Adult"
/dev_stage="Adult"
/organism="Homo sapiens"
/site_1: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      84 a      126 c      133 g      99 t
ORIGIN

```

```

Query Match      3.1%; Score 64.6; DB 9; Length 442;
Best Local Similarity 62.6%; Pred. No. 2.6e-07;
Matches 117; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY  922 cactggcaccatctcgacgcgtcaacggtatccatctcggtagtaagatccg-tgct 980
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  428 CATTGGGCATCATGATGCGGCGCAACTGTGTGCTGCTGCAAGATTGCGATTCT 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  981 gctgtccagcaagctctgtgtctgcacatctgtatgcagtcgacgaaaccttacc 1040
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  368 GCTGTCTCAGACGCTGCGCAGCCTTGTCGACCTCATACGTCGCCAGCAAAACGTTCA 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  1041 gttcccgagttgcgagcgctcttgatgagtcgacgaggaagatggttcttgagaa 1100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  308 TGTGGGCGAGCTGCGGCGCGTGTAGAGATGCAAGCAAGAACTGGATGCTTCGAA 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  1101 cgttaac 1107
    |||||
DB  248 CCTGAAC 242
    |||||

```

```

RESULT 2
BF962526      206 bp      mRNA      linear      EST 22-JAN-2001
LOCUS      BF962526
DEFINITION      PM1-NN1200-151200-013-h06 NN1200 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF962526
VERSION      BF962526.1 GI:12379801
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 206)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalhal,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.U., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.G.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&tl2=PM1-NN1200-
151200-013-h06&tl3=2000-12-15&tl4=1)
Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 206.
Location/Qualifiers

```

## FEATURES

```

FEATURES
Location/Qualifiers

```

```

source
1. 206
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN1200"
/dev_stage="Adult"
/organism="Homo sapiens"
/site_1: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      26 a      44 c      89 g      47 t
ORIGIN

```

```

Query Match      2.2%; Score 46; DB 10; Length 206;
Best Local Similarity 78.6%; Pred. No. 0.039;
Matches 55; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY  198 gccctgcggtctgcggtctgcggtctgcggtctgcggtctgcggtctgcggtctgt 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB  106 GCTAGTCGGGCGCTGCGGCGCTAGTCGGGCGCTAGTCGGGCGCTAGTCGGGCGCTAGT 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  258 cccgcctgtc 267
    |||||
DB  166 CGGGGCTGTC 175
    |||||

```

```

RESULT 3
BF960900      304 bp      mRNA      linear      EST 22-JAN-2001
LOCUS      BF960900
DEFINITION      MR3-NN0219-081200-012-f10 NN0219 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF960900
VERSION      BF960900.1 GI:12378175
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 304)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalhal,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.U., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.G.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&tl2=MR3-NN0219-
081200-012-f10&tl3=2000-12-08&tl4=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 303.
Location/Qualifiers

```

## FEATURES

```

FEATURES
source
1. 304
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN0219"
/dev_stage="Adult"
/organism="Homo sapiens"
/site_1: SmaI;

```







LOCUS	BF947371	312 bp	mRNA	linear	EST 22-JAN-2001
DEFINITION	MR3-NNO219-281000-006-g11 NN0219 Homo sapiens cDNA, mRNA sequence.				
ACCESSION	BF947371				
VERSION	BF947371.1	GI:12364646			
KEYWORDS	EST.				
ORGANISM	human.				
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 312)				
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare J.M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
MEDLINE	20202663				
COMMENT	Contact: Simpson A.J.C. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR3&c2=MR3-NNO219-281000-006-g11&t3=2000-10-28&t4=1) Seq primer: puc 18 forward High quality sequence stop: 59 High quality sequence start: 311. Location/Qualifiers 1. .312 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="NN0219" /dev_stage="Adult" /note="Organ: nervous normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."				
FEATURES	source				
BASE COUNT	53 a 72 c 107 g 80 t				
ORIGIN					
Query Match	2.1%; Score 44.4; DB 10; Length 312;				
Best Local Similarity	77.1%; Pred. No. 0.15;				
Matches 54; Conservative	0; Mismatches 16; Indels 0; Gaps 0;				
OY	198	gcctgctggaggtcttcaggatcggcggttgcggcgcttgtagggccgtccctttgt	257		
Db	203	GCTAGTCGGGCGGTGTCGGGCTAGTCGGGCTGTCGGGCTAGTCGGGCTGTCGGGCTACT	262		
OY	258	cgcgcctgtc	267		
Db	263	CGGGGCTGTC	272		
RESULT	7				
AM497369/c	392 bp mRNA linear EST 24-FEB-2000				
LOCUS	AM497369				
DEFINITION	9859a02.y1 Moss EST library PPU Physcomitrella patens cDNA clone				
PEP_SOURCE_ID	PPU_SOURCE_ID:PPU090903 5'', mRNA sequence.				
ACCESSION	AM497369				
VERSION	AM497369.1 GI:7067514				
KEYWORDS	EST.				

1	Source	Physcomitrella patens.
2	Organism	Physcomitrella patens
3		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
4		Brassicopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
5	Reference	1 (bases 1 to 392)
6	Authors	Quatrano,R., Bashardes,S., Cove,D., Cumling,A., Knight,C., Clifton
7		,S., Matra,M., Hillier,L., Pave,D., Martin,J., Wylie,T., Underwood
8		,K., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T.,
9		Stephoe,M., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.
10		,Waterston,R. and Wilson,R.
11		Leeds/Mash u Moss EST Project
12	Title	Unpublished (1999)
13	Journal	Contact: Ralph Quatrano
14	Comment	Leeds/Mash u Moss EST Project
15		Washington University School of Medicine
16		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
17		Tel: 314 286 1800
18		Fax: 314 286 1810
19		Email: est@watson.wustl.edu
20		Libraries were constructed by Dr. Stavros Bashardes as part of the
21		Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
22		Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
23		University Genome Sequencing Center For information on obtaining a
24		clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
25		Seq primer: 40RP from Gbpc
26		High quality sequence stop: 372.
27	Features	Location/Qualifiers
28	source	1..392
29		/organism="Physcomitrella patens"
30		/db_xref="taxon:3218
31		/clone="PEP_SOURCE_ID:PPU090903"
32		/clone_lib="Moss EST library PPU"
33		/tissue_type="Protonemata: 7 day old tissue
34		ammouni-grown"
35		/lab_host="DH10B"
36		/note="vector: pbluescript SK-; Site_1: EcoRI; Site_2:
37		XhoI; Construction of the cDNA library was carried out
38		using Stratagene's 'unizap - cDNA synthesis kit'. cDNA
39		was constructed using an oligo dt <sub>18</sub> primer/linker that
40		contains a XhoI site within it. Following ds cDNA
41		synthesis, EcoRI adapters were ligated to the blunt ends
42		and sample was digested with XhoI. The result is cDNA
43		with an EcoRI sticky end on one side and a XhoI sticky
44		end on the other. This cDNA was ligated directionally in
45		unizap arms. The vector is designed containing the
46		pbluescript sequence as well as lambda DNA and cDNA is
47		cloned within this pbluescript sequence. The vector was
48		then packaged using Gold gimpackaging extracts. Library
49		was grown in XLBlue MRF <sup>+</sup> cells and amplified. The library
50		was excised by mass excision using Stratagene's 'Mass
51		excision kit' that uses exasit as a helper phage that
52		releases the pbluescript sequence and circularises it as
53		single stranded plasmids that are then packaged (by helper
54		phage) and secreted out of the host cell as phagemids.
55		SOLR cells were transformed with phagemids and the library
56		was plated out on LB-amp plates to select for
57		transformants. Approximately 1,000,000 colonies were grown
58		and recovered. The double stranded plasmid library was
59		recovered by using Qiaagen Midi prep kit. 2 micro grams of
60		each library were used to transform DH10B cells by
61		electroporation."
62	Base Count	130 a 78 c 115 g 68 t 1 others
63	Origin	
64		Query Match 2.1%; Score 44; DB 9; Length 392;
65		Best Local Similarity 73.7%; Pred. No. 0.22; 20; Indels 0; Gaps 0;
66		Matches 56; Conservative 0; Mismatches
67	194	tttcgcctgcgcgcctgcgcctgtcgtcgcgcctgcgcgcctgcgcgcctgcgcctc 253
68	194	
69	217	ttccttcttgcgcgcctgcgcctgcgcctgcgcctgcgcctgcgcctgcgcctc 158



QY 254 ttgtccgcgcctgtcct 269  
11 11 11 11 11 11  
Db 157 TTATCGTCTGTCTCCT 142

RESULT 8  
BJ204205/c  
LOCUS  
DEFINITION BJ204205 497 bp mRNA linear EST 25-JAN-2002  
caulonemata and rhizoid-like protonemata Physcomitrella patens  
subsp. patens cDNA clone pph50b12 5', mRNA sequence.  
ACCESSION BJ204205  
VERSION BJ204205  
KEYWORDS GI:18372603  
SOURCE EST.  
ORGANISM Physcomitrella patens subsp. patens.  
Physcomitrella patens subsp. patens.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.  
REFERENCE 1 (bases 1 to 497)  
AUTHORS Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,  
Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe  
M.

TITLE Comparison of the moss Physcomitrella patens genome with flowering  
plants genome  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp

FEATURES  
source  
1. 497  
/organism="Physcomitrella patens subsp. patens"  
/db\_xref="taxon:145481"  
/clone="pph50b12"  
/clone\_1lb="normalized full length cDNA library,  
chloronemata, caulonemata and rhizoid-like protonemata"  
/tissue-type="mixture of chloronemata, caulonemata and  
rhizoid-like protonemata"  
BASE COUNT 149 a 105 c 150 g 93 t  
ORIGIN

Query Match 2.1% Score 44; DB 10; Length 497;  
Best Local Similarity 73.7%; Pred. No. 0.25;  
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 194 ttctgctgtcgcgcgtgtcgcgcgtgtcgcgcgtgtcgcgcgtgtcgcgcctc 253  
11  
Db 251 TCCCTCTGTGCGTGTCTCCTTCATCATGCTTCCTTATCGTCTGTGCTTC 192

QY 254 ttgtccgcgcctgtcct 269  
11  
Db 191 TTATCGTCTGTCTCCT 176

RESULT 9  
BJ183869/c  
LOCUS  
DEFINITION BJ183869 501 bp mRNA linear EST 24-JAN-2002  
caulonemata and malformed buds Physcomitrella patens subsp. patens  
cDNA clone pphb38n07 5', mRNA sequence.

caulonemata and malformed buds Physcomitrella patens subsp. patens  
cDNA clone pphb38n07 5', mRNA sequence.  
ACCESSION BJ183869  
VERSION BJ183869  
KEYWORDS GI:18351817  
SOURCE EST.  
ORGANISM Physcomitrella patens subsp. patens.  
Physcomitrella patens subsp. patens.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.  
REFERENCE 1 (bases 1 to 501)  
AUTHORS Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,  
Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe  
M.

TITLE Comparison of the moss Physcomitrella patens genome with flowering  
plants genome  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp

FEATURES  
source  
1. 501  
/organism="Physcomitrella patens subsp. patens"  
/db\_xref="taxon:145481"  
/clone="pphb38n07"  
/clone\_1lb="normalized full length cDNA library,  
chloronemata, caulonemata and malformed buds"  
/tissue-type="mixture of chloronemata, caulonemata and  
malformed buds"  
BASE COUNT 150 a 105 c 152 g 94 t  
ORIGIN

Query Match 2.1% Score 44; DB 10; Length 501;  
Best Local Similarity 73.7%; Pred. No. 0.25;  
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 194 ttctgctgtcgcgcgtgtcgcgcgtgtcgcgcgtgtcgcgcgtgtcgcgcctc 253  
11  
Db 251 TCCCTCTGTGCGTGTCTCCTTCATCATGCTTCCTTATCGTCTGTGCTTC 192

QY 254 ttgtccgcgcctgtcct 269  
11  
Db 191 TTATCGTCTGTCTCCT 176

RESULT 10  
BJ179016/c  
LOCUS  
DEFINITION BJ179016 546 bp mRNA linear EST 24-JAN-2002  
caulonemata and malformed buds Physcomitrella patens subsp. patens  
cDNA clone pphb24104 5', mRNA sequence.  
ACCESSION BJ179016  
VERSION BJ179016  
KEYWORDS GI:18346970  
SOURCE EST.  
ORGANISM Physcomitrella patens subsp. patens.  
Physcomitrella patens subsp. patens.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.



REFERENCE	1 (bases 1 to 546)			
AUTHORS	Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.			
TITLE	Comparison of the moss Physcomitrella patens genome with flowering plants genome			
JOURNAL	Unpublished (2002)			
COMMENT	Contact: Tadasu Shin-1 Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6836 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp			
FEATURES	A backbone of the vector is pBluescript II, that was in vivo excised from a modified I <sup>+</sup> S phase vector (Mo bi Tec, Germany). XhoI digested-5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested-3' end including poly-A tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000), Carninci, P. et al. Protonemata were blended by the POLYTROP, and then cultivated on the BCD medium containing 0.5um BA (benzylaminopurine) for 8 to 13 days under the continuous light.			
SOURCE	Location/Qualifiers 1..546 /organism="Physcomitrella patens subsp. patens" /db_xref="taxon:145481" /clone="npb24104" /clone_lib="normalized full length cDNA library, chloronemata, caulonemata and malformed buds" /tissue_type="mixture of chloronemata, caulonemata and malformed buds"			
BASE COUNT	163 a	116 c	163 g	104 t
ORIGIN				
Query Match	2.1%;	Score 44;	DB 10;	Length 546;
Best Local Similarity	73.7%;	Pred. No. 0.27;		
Matches	56;	Conservative	0;	Mismatches 20; Indels 0; Gaps 0;
Qy	194	ttctgccttcggagcttcggagcttcgcgagcttcgcgagcttcgcgagcttcgcgc	253	
Db	251	ttcctttttctgccttcgccttcgccttcgccttcgccttcgccttcgccttcgccttc	192	
Qy	254	ttgtccgcgcctgcctc	269	
Db	191	tttatcgcttcgttcct	176	
RESULT 11				
LOCUS	BJ197576/c	548 bp	linear	EST 24-JAN-2002
DEFINITION	BJ197576 normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata Physcomitrella patens subsp. patens cDNA clone pbn2b14 5', mRNA sequence.			
ACCESSION	BJ197576			
VERSION	BJ197576.1 GI:18365499			
SOURCE	EST.			
ORGANISM	Physcomitrella patens subsp. patens. Physcomitrella patens subsp. patens Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 548) Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T. , Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.			
TITLE	Comparison of the moss Physcomitrella patens genome with flowering plants genome			
JOURNAL	Unpublished (2002)			
COMMENT	Contact: Tadasu Shin-1 Center For Genetic Resource Information			

National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel.: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshni@genes.nig.ac.jp

A backbone of the vector is pBluescript II, that was in vivo excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI digested-5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested-3' end including poly-A tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000), Carninci, P. et al.

In Protenomata were blended by the POLYTRON, and then cultivated on the BCD medium containing 1mM NNA (naphthalene acetic acid) for 8 to 11 days under the continuous light.

Location/Qualifiers  
1..548  
/organism="Physcomitrella patens subsp. patens"  
/db\_xref="taxon:145481"  
/clone="phn2bl4"  
/clone\_1lb="normalized full length cDNA library,  
chloronemata, caulonemata and rhizoid-like protonemata"  
/tissue\_type="mixture of chloronemata, caulonemata and  
rhizoid-like protonemata"

BASE COUNT      163 a    116 c    165 g    104 t

ORIGIN

Query Match                  2.1%; Score 44; DB 10; Length 548;  
Best Local Similarity       73.7%; Pred. No. 0.27;  
Matches     56; Conservative    0; Mismatches    20; Indels    0; Gaps    0;

Oy    194    ttctgcgtgagcgttcgtcgaggcttgtcgagcttgtagggccgtatcccctc    253  
         ||| ||||| ||||||| ||||||| ||||||| ||||| ||||| ||  
Db    249    TCCTCCTGTCGTGGTGCCTTGCCACATCGAGCTTGCCTTTAFCGHGCTTGCTTC    190  
         ||| | || |||||

Oy    254    ttgtccgccgtgctct    269  
         ||| | || |||||

Db    189    TTATCTGCTTGTCTCT    174

RESULT 12  
BI033579                  217 bp       mRNA       linear       EST 14-JUN-2001  
LOCUS PM1-NN1200-190201-018-h03 NN1200 Homo sapiens CDNA, mRNA sequence.  
DEFINITION BI033579  
ACCESSION BI033579.1 GI:14440205  
VERSION EST.  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 217)  
Dias Neto, F., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsushima, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FADESP/LICR Human Cancer Genome



Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?pl1=PM1&ct2=PM1-NN1200-190201-018-h03&ct3=2001-02-19&ct4=1)

Seq primer: puc 18 forward  
High quality sequence start: 16  
High quality sequence stop: 217.  
Location/Qualifiers

## FEATURES

## source

1. 217  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN1200"  
/dev\_stage="Adult"  
/note="Organ: nervous,normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## BASE COUNT

38 a 39 c 101 g 39 t

Query Match 2.0%; Score 43.2; DB 10; Length 217;  
Best Local Similarity 68.2%; Pred. No. 0.26;  
Matches 60; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

OY 180 ggattactgaattctgctgtcggagctgtcggagctgtcggagctgtcggagctgtc 239

Db 126 gggtcaggagagagatgaggtgtcggagctgtcggagctgtcggagctgtc 185

OY 240 cgggcctgtcctctgtccgcgcgtgc 267

Db 186 cggcctactcggcgtgtcggcgtatgc 213

## RESULT 13

AZ675008/c

LOCUS 844 bp DNA linear GSS 14-DEC-2000  
DEFINITION Entamoeba histolytica Sheared DNA Entamoeba histolytica

ACCESION AZ675008

VERSION AZ675008.1 GI:11812154

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 844)  
Loftus,B., Van Aken,S. and Fraser,C.

AUTHORS Determination of clone end sequences from Entamoeba histolytica

TITLE HM1:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjoftus@tigr.org

Clones are derived from the Entamoeba histolytica HM1:IMSS sheared

DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 18

High quality sequence stop: 838.

Location/Qualifiers

1. 844  
/organism="Entamoeba histolytica"

/strain="HM1:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: PHOS1; Site\_1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of *E. histolytica* using a method described by Clark and Diamond (Clark, C.G. and Diamond, L.S. (1993) *Entamoeba histolytica*: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudo and B. Barel, Oxford University Press, 1993).

## BASE COUNT

146 a 177 c 84 g 437 t

Query Match 2.0%; Score 43.2; DB 12; Length 844;  
Best Local Similarity 71.2%; Pred. No. 0.59;  
Matches 57; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 367 ttaaaggagtaattggtccgcaaaaggagtaattggtccgcaaaaggagtaattg 426

Db 272 TCAAAAGGCAGAAATTGAAAGTCAAAGGCAGAAATTGAAAGTCAAAGGCAGAAATTGA 213

OY 427 gccgcaaaaggagtaattg 446

Db 212 AAGCAAAAGCAGAAATTG 193

## RESULT 14

CNS07CVP/c

LOCUS 1005 bp DNA linear GSS 08-JUL-2001  
DEFINITION T3 end of clone BD0AA007C03 of library BD0AA from strain CBS 94 of

Candida tropicalis, genomic survey sequence.

ACCESION AL439691

VERSION AL439691.1 GI:12223104

KEYWORDS GSS.

SOURCE Candida tropicalis.

ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; mitospotic Saccharomycetales; Candida.

REFERENCE 1 (bases 1 to 1005)  
Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and

Dujon,B. Genomic exploration of the hemiascomycetous yeasts: 16. Candida

tropicalis

JOURNAL FEBS Lett. 487 (1), 91-94 (2000)

MEDLINE 20584726

REFERENCE 2 (bases 1 to 1005)

Souciat,J.L., Aigle,M., Artiguenave,F., Blandin,G.,

Boitlin-Fukuhara,M., Bon,E., Brotlier,P., Casaregola,S.,

de-Montigny,J., Dujon,B., Duren,P., Leplinge,A., Lorente,B.,

Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,

Saurin,W., Tekala,F., Toffano-Nicche,C., Wesolowski-Douval,M.,

Wincker,P. and Weissenbach,J. Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

REFERENCE 3 (bases 1 to 1005)

Genoscope.

Direct Submission  
Submitted (08-SEP-2000) Genoscope - Centre National de Séquencage,

2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: [seqef@genoscope.cns.fr](mailto:seqef@genoscope.cns.fr) - Web: <http://www.genoscope.cns.fr>)

This GSS is part of a random genomic sequencing program of thirteen

yeast species: *Saccharomyces bayanus* var. *uvaurum*, *Saccharomyces*

*exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,

*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*

*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*

*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,

*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See

keywords for description of this sequence and for the sequence of







**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

ON nucleic - nucleic search, using sw model

Run on: August 20, 2002, 17:57:42 ; Search time 6893.08 Seconds  
(without alignments)  
17787.198 Million cell updates/sec

Title: US-09-826-206-3  
Perfect score: 5859  
Sequence: 1 tcgcgcgttcgcgtgac.....taccagagcccttcgc 5859

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_bt:\*  
3: gb\_in:\*  
4: gb\_lm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_inv:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
------------	-------------	-------	--------	-------	-------------

1	5859	100.0	5859	6	AX281575	AX281575 Sequence
2	2149.4	36.7	5108	6	AX138932	AX138932 Sequence
3	2149.4	36.7	5108	6	AX138933	AX138933 Sequence
4	2149.4	36.7	5108	6	BD000697	BD000697 Growth ho
5	2149.4	36.7	5108	6	BD000698	BD000698 Growth ho
6	2149.4	36.7	5111	6	AX138936	AX138936 Sequence
7	2149.4	36.7	5111	6	BD000701	BD000701 Growth ho
8	2149.4	36.7	5185	6	AX138928	AX138928 Sequence
9	2149.4	36.7	5185	6	BD000693	BD000693 Growth ho
10	2149.4	36.7	5188	6	AX138940	AX138940 Sequence
11	2149.4	36.7	5188	6	BD000705	BD000705 Growth ho
12	2149.4	36.7	5254	6	AX138941	AX138941 Sequence
13	2149.4	36.7	5254	6	BD000706	BD000706 Growth ho
14	2146.8	36.6	4779	6	AX080955	AX080955 Sequence
15	2146.8	36.6	4864	6	AX029455	AX029455 Sequence
16	2146.8	36.6	5310	6	AX080989	AX080989 Sequence
17	2146.8	36.6	5322	6	AX080956	AX080956 Sequence
18	2146.8	36.6	5900	6	AR064320	AR064320 Sequence
19	2146.8	36.6	5900	6	BD008816	BD008816 ComposIt1
20	2146.8	36.6	5952	6	AR064321	AR064321 Sequence
21	2146.8	36.6	5952	6	BD008817	BD008817 ComposIt1
22	2146.8	36.6	6709	6	AX080952	AX080952 Sequence
23	2146.8	36.6	6729	6	AX080954	AX080954 Sequence
24	2146.8	36.6	7099	6	AX180761	AX180761 Sequence
25	2146.8	36.6	7528	6	AX080953	AX080953 Sequence
26	2146.8	36.6	8366	6	AX180752	AX180752 Sequence
27	2145.2	36.6	5676	6	AR064322	AR064322 Sequence
28	2145.2	36.6	5676	6	BD008818	BD008818 ComposIt1
29	2145.2	36.6	5682	6	AR064323	AR064323 Sequence
30	2145.2	36.6	5682	6	BD008819	BD008819 ComposIt1
31	2145.2	36.6	7521	6	AX080951	AX080951 Sequence
32	2145.2	36.6	11358	6	BD008820	BD008820 ComposIt1
33	2123.2	36.2	4915	6	BD008822	BD008822 ComposIt1
34	2112	36.0	2112	6	AX281573	AX281573 Sequence
35	2007	34.3	2517	6	AX281576	AX281576 Sequence
36	2007	34.3	8509	6	AX281574	AX281574 Sequence
37	2007	34.3	8509	6	AX281868	AX281868 Sequence
38	1940	33.1	4296	6	I70200	I70200 Sequence 15
39	1940	33.1	4352	6	I70201	I70201 Sequence 16
40	1940	33.1	6047	6	I70197	I70197 Sequence 12
41	1940	33.1	6047	12	CYU40374	U40374 Luciferase
42	1928.4	32.9	4646	6	AX000085	AX000085 Sequence
43	1904	32.5	3993	6	I70199	I70199 Sequence 14
44	1904	32.5	6044	6	I70203	I70203 Sequence 18
45	1903	32.5	4120	6	I31863	I31863 Sequence 20

## ALIGNMENTS

RESULT 1  
AX281575  
LOCUS AX281575 5859 bp DNA linear PAT 02-NOV-2001  
DEFINITION Sequence 3 from Patent WO0177347.  
ACCESSION AX281575  
VERSION AX281575.1 GI:16608827  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
synthetic construct.  
artificial sequence.  
REFERENCE  
1 (sites)  
AUTHORS D'Elia, J.  
TITLE ketogulonigenium shuttle vectors  
JOURNAL Patent: WO 0177347-A 3 18-OCT-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
source  
1..5859  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="PADM291-4DS"

BASE COUNT	1444	a	1487	c	1566	g	1362	t
------------	------	---	------	---	------	---	------	---



Query Match	100.0%	Score 5859	DB 6	Length 5859
Best Local Similarity	100.0%	Pred. No. 0		
Matches 5859	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
QY 1	tcgcgcgttcctcgttctgacgtgtgcaaaacctctgacatgcagctccgcgaagcgtca	60		
Db 1	TCGGCGCTTTCGGGATGACGGGTGAAACCTCTGACATGCAGCTCCCGSAGACGGTCA	60		
QY 61	cagctgtctcgtlaagcggatgcgcggagacagaacaagccctcagggcgctcagcgggtg	120		
Db 61	CAGCTGTGCTGTAAAGCGGATGCGGAGACAGAACCCGTCAGGGCCGCGACGGGTG	120		
QY 121	ttgcgcgggtctgcgggcttgcttcaactatgcgcgcataagacagatgttactggaatgc	180		
Db 121	TTGGCGGGTGTGCGGGCTTGCTTAACATAGCGGCATCAGACGACATTTACTGAGAGTGC	180		
QY 181	accatatcgcgtgtggaatatccgcacacatctgcgtlaagagaaataaccgcatacagcgcc	240		
Db 181	ACCATATCGGGTGTGAATACCGACAGATGCTAGAGAAATACCGCATACGGCGCC	240		
QY 241	attgcgcattcaagcctgcgcgaactgtttggaaagggatcggctgcggcccttcgcata	300		
Db 241	ATTGCCCATTTAGGCTGCGCAACTGTTGGAAAGGAGATGGGTGGCGCTTTCGCTAT	300		
QY 301	tacgcgcacgtgcgaagaaggagatctgtctcgaaggcagatlaagtttggtaacgcgaaggt	360		
Db 301	TACGCCACGTGCGCAAGAGGGGAGATGCTGCMAAGGGGATTAAGTTGGTAAACGCACAGGT	360		
QY 361	tttcccgatcagcagcgttctgaataacgacgcgcagtgaaattcgcgaattggtctgaattca	420		
Db 361	TTTCCCGATCAGCAGCGTTGTAAACGACGGCGCATGTGAATTCGGCAATVGGGTGAAATTTCA	420		
QY 421	tagaatittgtgttgaggtgtgcgttagcgcgtctgcgaaggggtgcgcgcggaagatactctggt	480		
Db 421	TAGAAATTTGTGTGAGGTGTGCTAGCGGCTCTGACAGGGGTCGTGGCGGAGATCTCTGAT	480		
QY 481	ctcaaggtlaaggcgacaatctgagagaggttgaattgccccctgtatcgcctctcgtgtgcg	540		
Db 481	CTCAGGTAGGGCGCAATGAGAGAGGTGTAACTTCCCGCTGTCCCTCTCTGCGGGGG	540		
QY 541	catttggtatactctccgcggacataatgatatccgcctagaagatlaactgaatttctgc	600		
Db 541	CATTGGGTATCTCTGCCGACATATGATTTCCGCTAGAGAGATTAAGTTAATTCTCTGC	600		
QY 601	ctgtcgcgctgtctcgggctgtgcggcctgtgcggcctgtctgcggcctgtccctctgtcc	660		
Db 601	CTGTGCGGCTTGTGCGGGCTTGTGCGGGCTTGTGCGGGCTTGTGCCCTTGTCTCC	660		
QY 661	cgcctgtccctcaacttttcaacataaaaaaattggcgaaagccctctgttctcatagttc	720		
Db 661	CGCCTGTCCCTCACTTTTTCACATFAAAAAATGGCGCAAGCCCTTGTGTTATATGTTTC	720		
QY 721	ttatagttcaatacgaatatacacaataatcaatatacgtatctgccttaaaaggagta	780		
Db 721	TTATAGTTCAATACGAATAATTAACATTAATTAACATTAACCTTATTCGCTTAAAGGGAGTA	780		
QY 781	attgggcgcgaagaaggagtaatttggcgcgcaaaagggaatatttggcgcgcaaaaggga	840		
Db 781	ATTGGCGCGCAAAAGGAGTAATTTGGCGCGCAAAAGGAGTAATTTGGCGCGCAAAAGGA	840		
QY 841	gtaatttggccgatcgcgtgttcttcaatgggaggaatcccttaatcaattctccca	900		
Db 841	GTAATTTGGCGGATTCGGTGTGTTTCAATGGGAGGAATCCCTTAAATCATTTCTCCCA	900		
QY 901	tgggaagaacacacaagatggccgcagacgggctctgcacacagaacaaactgtgtcc	960		
Db 901	TGGGAAGACACAACAGTAGTGCGCGACCGGCTTTCACACAGCAAAAACCTGTGCTCC	960		
QY 961	ctgcgcaggtgtgcgaagggtctataatggaatccgcgcgcgcctgtcaagcgtctcaag	1020		
Db 961	CTGCCAGAGTGTGCAGAGGGGTCTATATGTGGCAATCCGCCCGCTCGACGGCGCTCAAGC	1020		

QY	1021	tcaatcatttaataagaccatctgcgggcgggcgcgaatgctctatgatgtgcgcataagaa	1080
Db	1021	TCATCATTTAATGATGATGCCACTCGCGGGCGGCCCATGAGCTGATGATGTCGCGCATGAA	1080
QY	1081	tgcgcgtgcgcgcacattgcgcgaatctgacggccatgaaaacccatgacgtggaagctga	1140
Db	1081	TCGCGCTGGCCGACATTTGGCCGAATCGACGGCATGAAAAACATGACCGTGAAGGCTGA	1140
QY	1141	ccccgctgtcgaagagctlaagccgctgtggtcttgaccatgatgacccctgcaaatga	1200
Db	1141	CCCCGCTGTTCAGGAGGCTAGCGCGCTGCGGTGTGACCCATGATGACCCCTGCAAAAGTGA	1200
QY	1201	tcgtacagtcggcggtcttgttctga tggagcgcaataagactacgcgcgaaggaagcgc	1260
Db	1201	TCGTGACATCTCGCGCTTGGTGTGATGAGCGCGCATGATGACTACCGCGCAGGAGCAAGCG	1260
QY	1261	ggaactcctagtgacgttgacccctccgagatlaattccgtctgataggcggagatctga	1320
Db	1261	GGAACCTCTTAGAGAGGTGAGACTTCCGAGATCTCCGCTGATGTGGCGCGGAGATCGA	1320
QY	1321	accactgggacattctcgacccgtcaaaoggtatctcatctcgtgataagatctcgtgc	1380
Db	1321	ACCACGTGGGCATTCTCGACCGTCAACCGATATCTCATCTCGATGATGATATTCGCGGC	1380
QY	1381	tgtctgttccagcgcgtctctag tctcccaactcttgatcgtatgtagcgcgaaaacctta	1440
Db	1381	TGCTGTTCCAGACAGTCTGATGCTCGCAACTTGATCGATGCGCGCGCAAAACCTTTA	1440
QY	1441	cggtcgcccgagattgvcgggcgtccctctgtagtgcccgagggaaaagatgtgtcgttgaagc	1500
Db	1441	CGGTCGCCGAGTTCGCGGCGCTCTTGTGAGTGCCTGAGGAAGATGCGTTGTTGGAAGC	1500
QY	1501	acgttaacgaattgctctcaaacctgcgacgtgatgagatacaacattatccgtctga	1560
Db	1501	ACGTTAACTGATTTGCTCTCAAACTCGCACTGGATGATCAACCATTTATTCGCGTCTGA	1560
QY	1561	catlgacgycgaagccgaaccaaatgtggccgtatgcgttgcaagtgttgactataagcttgg	1620
Db	1561	CATTGACGCGCAAGCCGACCAAGATTTGGCGGTAGCGTGGCAAGTGTGATATGAGCTTGG	1620
QY	1621	aaagtgaagagacccaacccgttcgcgaagcgcgcagatctgcgcggtttccaaagtctgaltcgag	1680
Db	1621	AAGTGAAGACGACGCCCAACCGTCCGACAGCGCGAGCGTGGGGGTTCCAAAGTCTGCGTCGAG	1680
QY	1681	atgtcgtctgcagaggggcagcggaaacgaatagccccctctctccagaagcggggcgga	1740
Db	1681	ATGCTCGTCGACGAGGGGCGACGCGGAACGATATCCCCCTCTTCCAGAAAGCGGGCGGA	1740
QY	1741	tcacctaaagtcacagtttggtcttgtagctctgaaacgctctgtctgcgcgcacaaaggaacag	1800
Db	1741	TCACCTACAGTCCACGTTGGGTGAGGTGAAGCGCTCTGCTGCGACCAACAAAGGACAAAGC	1800
QY	1801	atctgatcgctcgaactcgcgcgcttctgtcggagagaagcgtgctcttgagcgtg	1860
Db	1801	ATCTGATCGCTCAGACTTCCGGGCTTCTGTCTGCGGAGAAAGGCGTGCCTGTGGACGCTG	1860
QY	1861	caaacatcgaaaanaactgtttttagatctctgcgaaaagctagaaggaagtttgatttga	1920
Db	1861	CAAACTCTCAAAAACACTGTTTTAGATTTCTGCGCAAAAGTAAAGGAAGCTTTGAGTTTGA	1920
QY	1921	ggtaattcacgcgcaatagtgttaaatgacttcgttgaaacgaagtgcgaatatagcgttaa	1980
Db	1921	GGTATTTCAACCCCATAGTGTAAATGACTTTCGTGAACCATGTGCAATATGAGCGTTAA	1980
QY	1981	gactatgaaatacacagcgtctggaacagcctgcgaaaacgaacgcgggtgtggcgacgcgaaacat	2040
Db	1981	GACTATGAATAACACGGCTGCGACAGGCTGCAAAAGCAACGGGTGTGGCGAACGCATAT	2040
QY	2041	caactcgggcgtcaaaaagcgtgcaaatcttcggttaaaaaaga tgaatcttgggcata tgggt	2100
Db	2041	CACCTCGGGCGCTAAAAAGGCGTAAAAATTTCCGTTAAAAAAGATGAATCTGGGCGATGGGT	2100
QY	2101	tatagatcctgcggaattgcacagagtgttctctccattccaagaataacacccgaac	2160



|||||  
Db 2101 TATAGATCTCAGAAATTGCACAGAGTGTTCCTCCATTCCAAGAAATACCGAAAC 2160  
Qy 2161 aacctaaacgcaagratltgltgaagcgtgatagaaacacatgaatagacctgaagaatcag 2220  
Db 2161 AACTTAACGACGAATATATGTGAACGTGATGAACACATGAATAAGACCTCGAATATAG 2220  
Qy 2221 cgcattaaagcgtgaagtltcggaacttaacgcatctatctatgtatgcgaaggaagatcgt 2280  
Db 2221 GGCATTACAGGCGTGAAGTTTCGACTTTACGCATTCCTTTATCTGATGCCAGGAGAGATCG 2280  
Qy 2281 cgaacaaatgycgcgacatgycgcgagcgtcttcaattcaatccgatgtagagaggaaga 2340  
Db 2281 CGACAAATAGGCGCGACATGCGCGACGCTCTTCAATTTTCATCCGATGAGAGAGAGAAGA 2340  
Qy 2341 cgcgcgcctctaaacaaagaatgltggaagatatctgatacttcgttcgttcgaagcctt 2400  
Db 2341 CCGCCCCCTCAAAAACAAAGATGGTGAAGATTTGTATCTTGATCTGGCTTCAGGAGCCTT 2400  
Qy 2401 gccttactgycggaacaaacgcgatalttgaagcacagccgcacattagaagcgcgaagc 2460  
Db 2401 GCCTTTACTGGCGGAAAAACGGGATATTGAGGCAACAGCCCGCACTTTAGAGCGGAAAC 2460  
Qy 2461 ctataacagatlaacaaacactagaagccagaattgaaggaataaggaacgltggatcct 2520  
Db 2461 CTATTAACGATACCAAAACACTAGAACGACATGAGGAAATFAGGAACGTGGATCCT 2520  
Qy 2521 ctagaagtcgaacctgtagacgaatgcaagcttggcgtaatacgtgtcatagctgttccctgt 2580  
Db 2521 CTAGAGTGTGACCTGCAGGACATGCAAGCTTGCGTAAATCATGTCGTATGCTGTTCTCTGG 2580  
Qy 2581 tgaattgtlatcgcctcaacaattccacaacaatacagagccggaagcaataaaglttaa 2640  
Db 2581 TGAATTTGTATCCCTCTACAAATTCACACAAACATACGACCGGAAAGCATAAAGTGTAA 2640  
Qy 2641 gacctgggtgccttaatgtagagtaactacataatltgctgtgctgactgcccgt 2700  
Db 2641 GCCTGGGTGCTTAATGATGATGAGTAACTAACATTAATGTGCTGCTGCTACATGCCCCGT 2700  
Qy 2701 ttccagtcgggaacaccttcgtgccaagctgcatatgaatctgccaagccgcggggaga 2760  
Db 2701 TTCCAGTGGGGAACCTGTCGTCGACGTCATTAATGATAGGCGCAACGCCGGGGAAGA 2760  
Qy 2761 ggcgagtttcgctatctgggcgcctctccgcgcgcgttctgtccttgcgtcgtglatgta 2820  
Db 2761 GCGGTTTGGGTATTTGGGGGCTCTTCGCGGCTCGGTCTTGCTGCTGCTGCTGATATGA 2820  
Qy 2821 cttaacacgctcgcggaagtcgctcttcttgaatgagcgcatgagggaacgtgtcgtgcaat 2880  
Db 2821 CTTCAACGAGCTCCGGAAGTCCGCTCTTGTGATGAGGCGCATGGGAGCGTGCCTTGCAAT 2880  
Qy 2881 caagcgacccccccgcgcggttttgaagcgtataaagaatcatgctcctgcccgcggcgga 2940  
Db 2881 CACGGGCAACCCCCGCGCTTTTAAACGGCTTAAAAAAGTATGGCTCTGCTCGGGCGGA 2940  
Qy 2941 ccaagcccatcatgaacctgtccaagctcgtcctgtctctcctcgaatcttcgcgaacaggg 3000  
Db 2941 CCAGCCCATCATGATGACCTTGGCAACACTCGTCTGCTTCTTGATCTTCCGACAGAGGG 3000  
Qy 3001 cgaagatcgtgcatcatcaacgaacgcgcgcgtgagcgcggtcgtcgttgaagccgaaglttca 3060  
Db 3001 CGAGATGCTGGCATCACCGAACCGCGCGTGCGGGGGCTCGGAGACAGAGATTCA 3060  
Qy 3061 gcaagcgcccccaagcgccccaggttcgcaatgtatgcggcgcaagcttcgcgaacgtgctcat 3120  
Db 3061 GAGGCGCCCGCCAGGGCGCCCAAGTGCCTTATGATGCGGGCCAGCTGCGGAGCTGCTCAT 3120  
Qy 3121 agtccacagcgcccgltgatttctgtagccctgagcgacagcgccaggttagagccgaacgc 3180  
Db 3121 AGTCACAGACGCGCGTGAATTTGTAGCCCTTGCGACAGCGCACAGGTAGGCGCGACAGGC 3180  
Qy 3181 tcatgctggcgcgcgcgcttctcctcaatgcgtcttcgttcgttcgtgaagcagttaca 3240  
|||||

Db 3181 TCATGCCCGCCCGCCGCTTTTCCTCAATGCGCTTCCTGTTGCTGTGAAGCAGATACA 3240  
Qy 3241 ccttgaatagtgagcgttcgcttccttcgtgtgagctgtgttcatcaagccatcgcttgcct 3300  
Db 3241 CTTTATAGTGGGTGGCTGCCCTTCTGTGTGGTGGTTTTCATCAGCCATCCGCTTGCCCT 3300  
Qy 3301 catctgttaccgcgcgcggttagccggccagccttcgcagagcaagatltccggttgaacacg 3360  
Db 3301 CATCTGTTACGCCCGCGGAGTACCGCGACGCTTCGACAGACAGATATCCGTTGAGACCG 3360  
Qy 3361 ccaagltgcgaataaaggacagtgaagaagaacaccgcctgcgcggttgggcttactcac 3420  
Db 3361 CCAGGTGCGCAATTAAGGAGATGAAGAAACACCCGCTCCGCGGTGGCTTACTTTCAC 3420  
Qy 3421 ctatcctgcgcgcgtgagccggttgaatacaccagaagaaagcttacaagaaacctttgac 3480  
Db 3421 CTATCTGCGCCGGCTGACCCGTTGGATACACCAAGAAAGTCTACACCAACCTTTGGC 3480  
Qy 3481 aaaatcctgtatatacgttgcgaataaagatgataatacgaataaaatcgcataatgaccc 3540  
Db 3481 AAAATCCTGTATATGTGCGGAAAAAGATGATATACGAAAAAATCGCTATATAGACCC 3540  
Qy 3541 cgaagcaaggttatagcagcggaacagcttcctcgtcactgaactgcgtgcgtcgtgct 3600  
Db 3541 CGAAGCAGGTTATGACAGCGGAAAAAGCTTCTCGCTCACTGACTGCTCGCTCGGTCGT 3600  
Qy 3601 tgcgtctgcgcgcgagtgatltcagctcactcaagagggatataaggttatacagaatac 3660  
Db 3601 TCGGCTGCGCGGAGGAGATACAGCTACCTCAAAAGGGGATATACGTTATCCACAGAAATC 3660  
Qy 3661 aagggaatacgcaggaataaacaatgtgagcaaaagccagcaaaagccaggaacccgtaa 3720  
Db 3661 AGGGATTAACGAGGAAGAAACATGTGACCAAAAGGCCACGAAAAAGCCAGAACCGTAA 3720  
Qy 3721 aaagcgcggtgtcgtgagcttltccaatagctccgccccctgacagatcaacaataa 3780  
Db 3721 AAAGGCGCGGTGCTGCGCTTTTCCATAGCGTCCGCCCTCGAGACATCACAAAAA 3780  
Qy 3781 tgaagctcaagttcaagaggtggcgaaacccgcgaaggaactataaagatacagaagcgttcc 3840  
Db 3781 TCGAGCCTTAATTCAGAGGTGGCGAANCCGACAGAGACTTAAAGTATCCAGGCGTTTC 3840  
Qy 3841 cccctggaagcttcctcgtgctcctcgtlttcgaacctgcgcttaaccgatacctgtc 3900  
Db 3841 CCTGGAAACCTCCCTGTCGCGCTCTGCTTCGACCTTGCGCTTAACCGGATATCTGTC 3900  
Qy 3901 cgccttctcccttcgggaagcggtggcgttctcaatagctcaagctgtagatatacga 3960  
Db 3901 CGCCTTTCTCCCTTCGGAAGGTGGCGCTTCTCATFACCTCACGCTTAAGTATCTAG 3960  
Qy 3961 ttcggtgtaggtcgttcgcgttccaagcttgagcgtgtgacagaaaccccggttcaagccga 4020  
Db 3961 TTGCGGTAGGTGTTCCGTTCCAAAGCTGGGCTGTGTGCAAGAACCCCGGTTCAAGCCGA 4020  
Qy 4021 ccgctgcgccttataccgtaactatcgtcttgaagtcacaacccggtlaagaacagactatc 4080  
Db 4021 CCGCTGCGCTTATCCGTAACATATGCTTGAAGTCAACCCGTAAGAACAGACTATATC 4080  
Qy 4081 gccaactggaagcaagcaactgtgtlaacaggaattagcagaagcgaagltatgtagcggtgtac 4140  
Db 4081 GCCACTGCGCAGAGACCACTAGTAAACAGATTGACAAAGCAGAGGATATGAGCGGCTCTAC 4140  
Qy 4141 aagaattctgaaggttggtcctaactaagcgttaacataaagaagcgtatttggatctgc 4200  
Db 4141 AAGATTCTTGAAGTGGTGGCTTAACAGGCTAACCTTAAGAAAGACGATTTTGGTATCTG 4200  
Qy 4201 cgcctgtcgtgaagccagttactctcggaataaagagttgttagctcttgaatccgcgaacaa 4260  
Db 4201 CGCTTGCTGAAGCAAGTATACCTTGGAAAAAGATTGGTAGCTTTGATTCGGCAACAA 4260  
Qy 4261 aaccacacgctgtgtagcgtgttcttctgttgcgaagcagcagatatacgcgcaagaataa 4320  
Db 4261 AACCAACCGCTGTAGCGGTGTTTGTGTTTGTTCGCAAGCAGATTAACGCGCAGAAAAA 4320  
|||||







OY 3746 cataggtccgccccctgacgagcatcaaaaaatcgacgctcaagtcagagtgycga 3805  
|||||  
Db 724 CATAGGCTCCGGCCCCCTGACGAGCATCAAAAAATCGAGGCTCAAGTCAGAGGTGGCA 793  
OY 3806 aaccgcgaagactataaagatacagaaggtttcccccgtgaagctccctgtgctct 3865  
|||||  
Db 794 AACCCGACAGGACTATAAGATACAGAGGCTTCCCCCTGGAGAGCTCCCTCTGGCCTCT 853  
OY 3866 cctgttcgaacctgcgcttaacggatacgtctgcgcctttctcccttcggaagacgtg 3925  
|||||  
Db 854 CCTGTTCCGACCTTCGCCCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACCGTG 913  
OY 3926 gcgccttcataagtcacgctgtagtatacctcagttcgtgtgtagtgcgtccgaag 3985  
|||||  
Db 914 GCGCTTCTCATAGCTACAGCTGTAGGTATCTCAGTTGGTGTAGGTGCTTGCTGCACAG 973  
OY 3966 ctggagctgtgacgaacccccgttcaagccgacgctgcgcttalcggttaactat 4045  
|||||  
Db 974 CTGGGCTGTGTCCAGAACCCCCCTTCAGCCCGAACCGCTGCGCTTATCCGGTAACGTAT 1033  
OY 4046 cgtcttgagtcgaaccggttaagacagactatacgcacgttgacgacgacgacgttaac 4105  
|||||  
Db 1034 CGTCTTGAGTCCACCCGGTAAAGACACACTTATGCCACTGGACACACCTGGTAAAC 1093  
OY 4106 aggattacagagcgaaggtatgttagcggtgctacagagttcttgaaagtgcgttaac 4165  
|||||  
Db 1094 AGGATTAGACAGACGAGGTATGTAGCGGCTGTACAGAGTTCTTGAAGTGGTGGCTAAC 1153  
OY 4166 taagcttacactaagaagaagatttggtaactcgtgcgtctgcgtgaagccagttacctc 4225  
|||||  
Db 1154 TAGGCTACACTAGAAAGAACATATTGTATCTCGCTCTCTGAAGCACTTTACCTTTC 1213  
OY 4226 ggaataaagattggttagctcttcgtaccgcaaaacaaacacgcgtgcgtgagtgcttt 4285  
|||||  
Db 1214 GGAATAAAGATTGTTAGCTTGTGATCCGGCAACAAACACCGCTGTGTAGCGGTGTTT 1273  
OY 4286 ttgtttgcaagcagacgaattacgcgcgcaaaaaaagaatcctaagaagatccttgatc 4345  
|||||  
Db 1274 TTTGTTTGCACAGCAGATATACGGCAGAAAAAAGATCTCAAGAAAGATCCTTGTATC 1333  
OY 4346 tttcttaaggggtctgagcgtcagtggtgaacgaatacgcgttaagggttttgcgtcatg 4405  
|||||  
Db 1334 TTTTCTACGGGCTGTACGCTCAGTGAAGCAAAACCTCAGTTAAGGATTTTGGTCATG 1393  
OY 4406 agattatcaaaaagatctcaccatagatccttt----- 4440  
|||||  
Db 1394 AGATTATCAAAAAGATCTTCACTAGATCTTTTAATTAATAAATGAAGTTTAATCA 1453  
OY 4441 ----- 4440  
Db 1454 ATCTAAAGTATATAGTAAACTGTGCTGACAGTTACCAATGCTTATCAGTGAAGCA 1513  
OY 4441 -----ggggggggggggg 4452  
Db 1514 CCTATCTACGATGTGCTATTTGCTTCATCCATAGTTGCTGACTCGGGGGGGGGGG 1573  
OY 4453 cgcctaggtctgcctcgtaagaaggtgtgtctgactatacagaacgcgtaactgcgccca 4512  
|||||  
Db 1574 CGCTAGAGTCTCCTCGTGAAGAAGGTGTGCTGACTCATACAGGCTCGAATCGCCCA 1633  
OY 4513 tcaaccaagcagaagaatgagggagccaaggttgatgagaacgttgttgaagtgtgacag 4572  
|||||  
Db 1634 TCATTCAGGCAAGAAAGTGGAGCCACGCTGATGAGAGCTTGTGTAGGTGACACAG 1693  
OY 4573 ttgggtgatttgaactttgcttgcgcaaggaacggttcgctgtgcggaagaatgctgtg 4632  
|||||  
Db 1694 TTGGGTATTTTAACTTTTGTTCCTTGGCACGGAAGGCTGTGCTGTGCGGAAGATCGTG 1753  
OY 4633 atctatccttcaactcagcaaaagttcgaattatccaacaaagccgcgtccgcgtcaag 4692  
|||||  
Db 1754 ATCTGATCTTCACTACACCAAAAGTTCGATTATTCACCAAAAGCGCGCTCCCGTCAG 1813  
OY 4693 tcagcgtaattgtctgtgcaggtgttacaaccaatlaaccaatcttgatlaagaanaactat 4752  
|||||

Db 1814 TCAGGCTAATGCTCTGCCAGTGTACAAACATTAACCAATTAATCTGATTGAAAAACTCAT 1873  
OY 4753 cgaactcaaatgaaactcaattatcatatcaagattatcaatatacatattttgaa 4812  
|||||  
Db 1874 CGAGCATCAAAATGAACCTCAATTTATTCATATACAGGATTAATCAATATTTTGA 1933  
OY 4813 aaagccgttctgttaagagaagaactcaacgcggagcgttcaatagaatgagaaat 4872  
|||||  
Db 1934 AAAGCGTTTCTGTATTAAGAGAGAAACTCACCGGACAGTTCCATAGATGGCAAGAT 1993  
OY 4873 cctgtatcgtctgcgtatccgactcgtcccaacatcaatatacatatttccct 4932  
|||||  
Db 1994 CCGTGTATGCTGTGAGTATTCGACTCGTGCACATCAATCAACTATTAATTTCCCT 2053  
OY 4933 cgtcaaaaataaggtatacaagtgaagaatcaacatgagtgaagcactgaatccggtgaga 4992  
|||||  
Db 2054 CGTCAAAAATTAAGTTTATCAAGTGAAGAAATCACCAATGAGTGAACGTGAATCGGTGA 2113  
OY 4993 atggcaaaagctatgcatcttctccagactgttcaacagccagacatcagctcgt 5052  
|||||  
Db 2114 ATGCAAAAGCTTATCATTTCTTCCAGACTTGTTCACAGGCCAGCCATTACGCTCGT 2173  
OY 5053 catcaaaatcactcgtacatcaaccaaacgltatcatctgtgaltgycgctgagcgagac 5112  
|||||  
Db 2174 CATCAAAATCAGTCGATCAACCAACCGTTATTCATTCGTGATTCGCTGAGCGAGAC 2233  
OY 5113 gaaatacgcgactgcgtgttaaaagacaaatttaacaaaggaatcgaaatgcagccgca 5172  
|||||  
Db 2234 GAAATAGCGGATCGCTTTAAAGGACAAATTAACAAAGGAATCGAATCGAACCGCGCA 2293  
OY 5173 ggaacactgcgcgcgcatcaacaatacttcaacctgaatcaagaatacttcttaact 5232  
|||||  
Db 2294 GGAACACTCGCAGCGCATCAACATATTTTCACTGATTAATGAGATTTCTCTAATACCT 2253  
OY 5233 ggaatgcttctccgggagtcgcagtggtgaaatcaatgacatgacatcaagaatgacga 5292  
|||||  
Db 2354 GGAATCTGTTTCCCGGGATGCGAGTGTGATTAACATGATCATGAGTACGAGTACGGA 2413  
OY 5293 taaagtctgtatgctcggaagaagacataatccgtccagccagtttaactcatatc 5352  
|||||  
Db 2414 TAAATGCTTGAATGTCGGAAGGACATTAATTCCTCAGCCAGTTTATCTACCATCT 2473  
OY 5353 catctgaacatcatctggcaacgctacacttgcacgttccaagaacacactgcgcacat 5412  
|||||  
Db 2474 CATCTTAACATCATTTGGCAACGCTTACCTTGGCATGTTTCAGAAACACTCTGGCGCAT 2533  
OY 5413 cgggcttcccatcaaatcgaatagatgtgcgacactgattgcggaatatacgcgaccc 5472  
|||||  
Db 2534 CGGGCTTCCCATTCATTCGATAGATGTGCGACCTGATTGCGGACATTTATCGGAGCC 2593  
OY 5473 attataccatataaatacgaatccatgttgaatttaacgcgcgccttcgaagaag 5532  
|||||  
Db 2594 ATTTATACCATTAATTAATGACATTCATGTGAATTTAATCCGGCCTTCGACCAAGACG 2653  
OY 5533 ttcccgctggaatagtcacatacaaccccttataactcgttcaatgaagcgaagat 5592  
|||||  
Db 2654 TTTCCGTTGAATATGCTCATATACACCCCTGTATTACGTTTATGTAAGCAGACAGTT 2713  
OY 5593 ctatgtlcatatgatatatacttcatctgtgcaatgtaacaatcagagaatttgagaa 5652  
|||||  
Db 2714 TTTATGTTATGATGATTAATATTTTATCTGTGCAATGTAACATCAAGAGATTTTGAACA 2773  
OY 5653 caacgtgcttcccccaccccccataatgaagcattatacgggttatgtgctatga 5712  
|||||  
Db 2774 CAAGTGGCTTCCCCCCCCCCCATTTATGAAGCAATTTATCAGGGTATTGTCTCATGA 2833  
OY 5713 gcggaatacatattgaaatgatttaataaataaataaagaaggttcgcgcacattc 5772  
|||||  
Db 2834 GCGGATACATATTTGATGATTTTGAAGAAATTAACAAATAGGGGTTCCGCGCACATTTTC 2893  
OY 5773 cccgaaagtgcacactgacgttcaagaacaaatctatcatatgacattaaactataaa 5832  
|||||



[illegible][illegible]



QY	5113	gaataacgcgaatcgcgtgttaaaagagacaattacaacacggaatcgaatgcaacccgcgcga	5172
Db	2234	GAATAACCGCATCGCTGTTAAAGAGCAATTATCAAAACAGGAATCGAATCGAACCGCGCA	2293
QY	5173	ggaacacgcgaagcgatcaacaatatatttaacctgaatcagatattcttctaact	5232
Db	2294	GGAAACACGCGACCGCATCAACAATATTTTCACTGAATCAGGAATATTCCTTCTAATACCT	2353
QY	5233	ggaatgcgttttcccgaggatcgcagtggtgagtaacatcgaatcgaagagtaagca	5292
Db	2354	GGAAATGCTGTTTTCOCGGGGATCCGAGTGGTGAAGTAACCATGATCATCAGAGTACGGA	2413
QY	5293	taaaatgcttgaatcgatcggaagagcacaatctccatgcagcaagttagtcgaacct	5352
Db	2414	TAAATGCTTATGATGCTCGGAAGAGCATAAATTCCTGTCAGCCAGTTTGTCTGACCACT	2473
QY	5353	catctgtaacgaatcttgacaagcctactcttgcaatgctttagaanaacacctcagcat	5412
Db	2474	CATCTGTACATCATTTGGCAACGCTACCTTTGGCCATGTTTACAGAAACAACCTGGCGAT	2533
QY	5413	cgagcttcccatatacatgatatgctgcgaacctgatactgcgcgacatatacgcgaagcc	5472
Db	2534	CGGGTCTCCCATPACAACTGATAGATTGTGCGACCTGATTTGCCGACATTAATGCGCAGCC	2593
QY	5473	attataccatataaatacgaatccatcgaatgcttggaattaatcgcgctcgagcaagcg	5532
Db	2594	ATTATACCATATPMAATCACATCCATCGATTGGAATTTAAACGCGCCTCGAGCAAGACG	2653
QY	5533	tttcccgctgaataggctgaacaccccttgatctctttagtaagcagaacct	5592
Db	2654	TTTCCCGTGAATAGGCTCAATACACCCCTGTGATTAAGTTATGTATGTAAAGACACATTT	2713
QY	5593	ttatgttcatagaatataattctatcttgcgaatgaaacatcagaagatcttgaaga	5652
Db	2714	TTATGTTCATGATGATATATTTTATCTTGGAATGTAAACATCAGAGATTTTGAGACA	2773
QY	5653	caacgtgacctcccccccccccatattgaagcattatcgaaggtatattgtctcatga	5712
Db	2774	CAACTGTGCTTTCCCGCCCCCATTTATGAAAGATTATACAGGTTATTTGTCTCAAGA	2833
QY	5713	gcggaatacatattgaatgtaattagaanaataacaataaggggttcgcgcgaaccttc	5772
Db	2834	GCGGATACATATTGTAATGTATTATGAAATAATAACAATATAGGGTTCGCGCACATTTTC	2893
QY	5773	cccgaaagatgcacactgaagcgtctaagaacacatattatcatgacatlaaacctataaa	5832
Db	2894	CCCGAAAGATGCCACCTGACGCTPAAAGAACCATATTATCATGACATTAACTATAAA	2953
QY	5833	atagcgtatacagaagcccttcgtc	5859
Db	2954	ATAGGCGTATCACGAGGCCCTTTCGTG	2980
RESULT 4			
LOCUS	BD000697	5108 bp	DNA linear
DEFINITION	Growth hormone and growth hormone-releasing hormone composition.		
ACCESSION	BD000697		
VERSION	BD000697.1	GI:1862810	
KEYWORDS	JP 2000350590-A/50.		
SOURCE	synthetic construct.		
ORGANISM	synthetic construct.		
REFERENCE	1 (bases 1 to 5108)		
AUTHORS	Morsey,M.A. and George,M.		
TITLE	Growth hormone and growth hormone-releasing hormone composition		
JOURNAL	Patent: JP 2000350590-A 50 19-DEC-2000;		
COMMENT	PFIZER PROD INC		
	OS Artificial Sequence		
	PN JP 2000350590-A/50		
	PD 19-DEC-2000		
	PF 12-APR-2000 JP 2000111120		

[illegible]











QY	5473	attatatacccatataaatacagatccatcagtgttgaaatttaacatgcgcgccttcagagcaagc	5532
Db	2594	ATTATATACCCATTATTAATACGATCCATCGATGTTGGAAATTTAATCCGGCGCTCGAGCAAAACG	2653
QY	5533	ttccccgttgaaatgtgctcaaacacccttgattacgtctttatcattgaacagacagtt	5592
Db	2654	TTTCCCGTTGAATATGGCTCTAAACACCCCTTGATTAAGTATTATGTATGAAGCAACAGTT	2713
QY	5593	ttatgttcacgatgatataattttatctgtgcgaatgaacatcagagattttgagaca	5652
Db	2714	TTATGTCTCANGATGATATATTTATCTGTGCAATGTAACTACATCAGAGATTTTGAGCA	2773
QY	5653	caagtgaccttcccccccccccatatttgaagcatatcagagttatgtctcatga	5712
Db	2774	CAACGTGGCTTTCCCCCCCCCATATATGAGCATTTATCAGAGGTTATGTCTCAAGA	2833
QY	5713	gcggaatacatattgaatgtatttgaaaaaataaacaataaggggttcgcgcacatttc	5772
Db	2834	GGGGATACATATTTGAATGTATTTTGAAATAATAACAATAATAGGGGTTCCGGGCACATTTTC	2893
QY	5773	cccgaanaatgaccacctgaagctctaagaacaacattatataatcagatctaaccctataaa	5832
Db	2894	CCCGAAAATGGCCACCTGCAGTCTAAGAAACCATTTATATCATGTGACATTAACTATATAAA	2953
QY	5833	ataaggatatacgaagggcccttcgtc	5859
Db	2954	ATAGGCGTATCAAGAGGCCCTTTTCGTc	2980

[illegible]

OY	3806	aaccgcgaagactataaagataccaggcggttcccccgtggaagctccctcgtgagctct	3865
Db	794	AACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCTCGAAGACTCCTCGTGCCTCT	853
OY	3866	cctgtctccgaacctccgcttaccggataacctgtccgccttcccttccttcgggaagcgtg	3925
Db	854	CCTGTTCCGACCCCTCGCCTTACCGGATACCTGTCGCCCTTTCCTCGGGGAAGGTG	913
OY	3926	gcgccttctcaagctcaagcgtgtagagatcagatcagtcggtgtgtagtgcgttcgtccaaag	3985
Db	914	GGCGTTTCATATGCTCAAGCTCAAGCTGATGATCTCAATTGGGTAGGTGTCGTCCAG	973
OY	3986	ctggagctgtgtcacaagaaacccccgtltaagcccgacgcgtgcgccttaccgtaactat	4045
Db	974	CTGGGCTGTGTGACAGAACCCCCCTTACGCCCGACCCCTCGCCTTATCCGGTAACAT	1033
OY	4046	cgctcttgatccaaccccggtlaagacagactatcgcacatcggcagcagcacaactgttaac	4105
Db	1034	CGCTTTGAATCCAAACCCGGTAAAGACACACTTATGCGCACTGCGACGACCACTGGTAAAC	1093
OY	4106	agattatgacagagcgaagatgtlaagcgggtgtacagagttcttgaagtgtgtgcctaac	4165
Db	1094	AGGATTATGACAGACCGAGGATATGAGGCGGCTGTACAGATTCCTTAAGTGGTGGCTTAAC	1153
OY	4166	tacggtctaacctagaagagcagatatttggtatctcgctcgtctgtcaagccagttaaccttc	4225
Db	1154	TACGGCTACACTGAAGAAACAGTATTGGTATCTGGCTCTGCTAAGCAATTAACCTTC	1213
OY	4226	ggaanaagatltgtagctcttctgtatccggcaaaacaacacgcgtgtgtagcgtgtgtttt	4285
Db	1214	GGAAAAAGATGTGTGATGCTCTTATGTCGGCAACAAACACCGCTGTGAGCGGTGTTT	1273
OY	4286	tttgtttgcaagacagcagatctacgcgcagaanaaaagatctccaagaagatcctttgac	4345
Db	1274	TTTGTTTGCAACGACGACGAGATTACGGCGAAGAAAAAGATCTCAAGAAATCTTTGATAC	1333
OY	4346	tttctcaagggtctgacgctcagtgtagaacgaaaactccgttaagggatttgggtcatg	4405
Db	1334	TTTTCTCTACGGGCTCTGACGCTCAGTGAAGCAAAACCTCACTTAAGGATTTTGGTCATG	1393
OY	4406	agattatcaaaaagatccttcacatagatccttt-----	4440
Db	1394	AGATTATCAAAAAGATCTTACACTAGATCCTTTAAATTAAAAATGAATTTTAAATCA	1453
OY	4441	-----	4440
Db	1454	ATCTAAAGTATATATGAGTAACTTGGTCTGCACAGTTACCAATGCTTAATCACTGAGGCA	1513
OY	4441	-----gggggggggggg-----	4452
Db	1514	CCTATCTGAGGAGACTGCTCTCTTTTCGTTTCATCCATAGTTCCTGACACGGGGGGGGGGG	1573
OY	4453	cgcgtgaagctcgtcgtcggaagaaggtgtgtcttaaccatacacagagcgttaacgcacca	4512
Db	1574	CGCTGAGGCTGTGCTCGTGAAGAAAGGTGTTCTTACATCAATACAGGGCTCAATCGCCCCA	1633
OY	4513	tcatccagccagaagatgagggagccacggtltgtagagcttgtgtgtgagtgagacag	4572
Db	1634	TCATCCAGCCAGAAGTAGAGGAGCCACAGGTGTGATGAGAGCTTTGTTGTAGGTGACACAG	1693
OY	4573	ttgtgtgatttgaacttctgtctcttgcctccagcggaaagcgtctgctgtctcgggaagatcggtg	4632
Db	1694	TTGGTGATTTTAACTTTTGTCTTTCCTCCACGGAACGCTGTGGCTTGGGGAAGATGCGTG	1753
OY	4633	atctgatcttcaactcagcaaaagttcgattatcaacaagcgcgcgtcccgctaag	4692
Db	1754	ATCTGATCCTTCACATCAGCAAAAGTTGATTTATTCAAAAGCGCGCTGCTCCGCTCAAG	1813
OY	4693	tcaagctaatgtcttgcgcaggtgttcaaaccaatlaaccaatcttgatattagaanaaactcat	4752
Db	1814	TCAGCGTAATAGCTCTGCCAGTGTATCAACCAATTCGTATTTAGAAAAATCAT	1873



QY	4753	cgagcatccaatgaaactgtaatttatlctaatatcaggaattatcaatccatatltttgaa	4812
Db	1874	CGAGCAATCAATGAAATGCAATTTTATCATATCAGAGATTATCAATACCATTTTGTGA	1933
QY	4813	aaagcgtttctctlaaagaaggaanaactcaaccggaggaagtctcatgagtggcaaat	4872
Db	1934	AAAGCCCTTTCTGTATGAAAGGAAANACTCACCGAGGAGTTCATGAGATGGCAAGAT	1993
QY	4873	ccctgtaicgtctcgtcagatctcgaactcgtctcaacatcaataaacctatlaattccccc	4932
Db	1994	CCCTGTATGCTGGTCCGATTCGCACTCGTCCAAACATCATACACCTTTATTTCCCT	2053
QY	4933	cgctaaataaaggatltcaagttagaanaatcaccaatlgagtgagacgtgaatccgltga	4992
Db	2054	CGTCAAAAATTAAGGTTTATCAAGTGAAGAAATTCACATGATGTACGACTGAATCCGCTGAGA	2113
QY	4993	atggcaaaagcttatgcattcttctccagaactgttcaacaaagccagccatlaagctgt	5052
Db	2114	ATGGCAAAAGCTTATGATATTCTTTCAGACTGTGTCAACAGGCCAGCCATTACCTGT	2173
QY	5053	catcaaaatcactclogcatcaaccaaaacgltlctcaatltcgatltgcgcgtgaagagag	5112
Db	2174	CATCAAAATCATCTGCATCAACCAACCGTTATTCTATCTGTTATTCGGCTGAGGAGAC	2233
QY	5113	gaaatcagcgatcgcgtgttaaaagcaaatlatacaacaggaatcgaatgcaacgcgagca	5172
Db	2234	GAAATACGCGATCGCTGTTTAAAGAGCAATTTCAAAACAGAAATCAAAATGCAATCGCGCA	2293
QY	5173	ggaacctgcagagcatcaacaaatlttcaacctgaatcgaatattctctaact	5232
Db	2294	GGAACACTGCGCAGCGCAATCAATTTTTCACCTGAATCAGAAATTTCTCTAATACCT	2353
QY	5233	ggaatgctgttttcccgaggatcgcaagtgtgagtacaatgcatcaatcaaggaatagga	5292
Db	2354	GGAATGCTGTTTCCCGGGAATCGCACTGTGATGATACCTCATCATCAGAGATACGGA	2413
QY	5293	taaaatgcttgatgtlctggaagaggaataatltcgtlccagccaaglttgaactgaact	5352
Db	2414	TAAATGCTTGATGCTGGAAGAGGATTAATTCGTCAGCCAGTTTATGTGACATCT	2473
QY	5353	catctgtaaacatcatltyggcaagcgtacaccttgcattglttcaagaaacaactctggcgcat	5412
Db	2474	CATCTGTAAACATCATTTGGCAACCTCATCTTGGCATGTTTCAGAAACAACCTCTGGCGCAT	2533
QY	5413	cgaggttcccatlaaactcgaatagatgttgcacactggaattggccgagcatatctcgagagcc	5472
Db	2534	CGGCTTCCCATACATATCGATTAATTTGTGCACCTGATTTGCCGACATTAATGCGAGGCC	2593
QY	5473	attatacccatlaaatacgaatltcaatgtltgaaattlaatcggcgacctcgacaagaagc	5532
Db	2594	ATTATATCCCATTAATTAATTCAGCATCCATGTGTGAATTTAATGCGGCGCTCGAGCAAGCG	2653
QY	5533	tttccggttgaatltggtcctataaacacctgttatctactgtttatgtaagcagaagtt	5592
Db	2654	TTTTCCCTGTGAATATGGCTCATTAACACCCCTGTGTATACGTGTTTATGTAAGCAGCACTT	2713
QY	5593	ttatctgtcatgaatgaataatlttctactctgtgcaatgtaacatcagagatlttgaaga	5652
Db	2714	TTATGTTTCATGATGATATATTTTATCTTTGTGTGCAATGTAACATCAGAGATTTTGAACA	2773
QY	5653	caacgttgcttcccccccccccatlatagaagcatlatacaggttaltgtlctcatga	5712
Db	2774	CAACGTGCTTTCCCCCCCCCCCATTAATTAAGAGCATTTATACAGGGTTATTTGTCTCANGA	2833
QY	5713	ggcgatataattgaaatgtatcttagaaaaataacaataaggggttcggcgacatttc	5772
Db	2834	GGCGATACATATTTTGATTTGATTTTATAGAAAAATTAACAAATATAGGGGTTCCGCCACATATTTC	2893
QY	5773	ccgaaagaatgcccctcgagaagctctaaagaacattatctctgcattaaacctataaa	5833
Db	2894	CCCGAAAGATGCCACCTGAGGCTTAAGAAACCATTTATTTATCATGACATTTAACTTAATAAA	2953
QY	5833	ataggcgtlatacgaagcccttccgtc	5859

[illegible]



QY 3986 ctggagctgtgtgacgaaccccccttcagccgacgcgtgccttaccgtaactat 4045  
Db 974 CTGGGCTGTGTGACGAACCCCCGTTGACGCCGACCGCTGCGCTTATTCGGTAACATAT 1033  
QY 4046 cgtcttgatccaaccggtlaagacaacgaactatcgccaactgagcagccaactgtlaac 4105  
Db 1034 CGTCTTGAGTCCACCCGCTAAAGACAGACTTATCGCACACTGCGACGACGCTACGTGTAC 1093  
QY 4106 aggaattagcagagcgaggtatgtagcggtgtctacagagttctggaagtgtggttaac 4165  
Db 1094 AGGATTAGCAGAGCGAGGTATGTAGCGGTCTACAGAGTCTTGAGTGTGGCTTAC 1153  
QY 4166 tacggttaactagagaagcaagatttgtatctgcgtctgcgttgaagccagttaccttc 4225  
Db 1154 TACGGCTACACTAGAAAGAAAGATTTTGTATCTGCGCTGCTGAAACCCAGTTTACCTTC 1213  
QY 4226 ggaataaagagttgttagctcttgatccggtcaacaacacgcgtgtagcggtgttt 4285  
Db 1214 GGAATAAGAGTTGTAGCTCTTGATCCGGCAAAACACCGCTGCTGAGCGTGGTTT 1273  
QY 4286 ttgtttgcaagcagcagattacgcgcagaaaaaagagatcacaagaagatccttgaac 4345  
Db 1274 TTGTTTGCAGACGACAGATTACGCCACAAAAAGATCTCAGAGAGATCTTTGATC 1333  
QY 4346 ttcttaacggtgtcgtgacgcctagtggaacgaataactcagttlaaggttttgcctg 4405  
Db 1334 TTTTTCAGGGGCTGACGCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTATG 1393  
QY 4406 agatttcaaaaagagatcttcacactagatccctt----- 4440  
Db 1394 AGATTATCAAAAAGATCTTACCTGATACCTTTTAAATGAAGTTTAAATCA 1453  
QY 4441 ----- 4440  
Db 1454 ATCTAAGTATATGAGTAAACTTGTGTGACGTTACCAATGCTTAAATCAGTGAAGCA 1513  
QY 4441 -----999999999999 4452  
Db 1514 CCTATCTCAGGAGCTGTCTTATTTCCGTTTCATCCATAGTTGCCGACGCGGGGGGGGG 1573  
QY 4453 cgcctgagctctgctcgtgaaagaagtgtgtcgtgactacatacagagcctgaaatcgccca 4512  
Db 1574 CGGTGAGGTGCTGCTCGTGAAGAAGGTGTGCTGACTCAATACAGCGCTGAATGCGCCCA 1633  
QY 4513 tcaaccagccgaagaagtgaggagccaaggttgaatgagagcttgttgaatgagacag 4572  
Db 1634 TCATCCAGCCAGAAAGTGAAGGAGCCACAGGTGTGATGAGACTTGTGTGAGTGAACAG 1693  
QY 4573 ttgtgtgacttgaactttgtcttgcacaggaacggtctgcgtgtgcgggaagaatgtgtg 4632  
Db 1694 TTGGGTGATTTTGAACCTTGTCTTGGCCACGGMACGGTCTGCGTGTGGGAAAGTGGTG 1753  
QY 4633 atctgatacctccaactcagcaaaagttcgatttcaacaagaacgcgcgtcccgtaag 4692  
Db 1754 ATCTATATCTTCACTCAGCAAAAGTTGATTTATTCACAAACCCGCGTCCGTCAG 1813  
QY 4693 tcaagcgtatgctctgcagtggttacaacccaatlaacaaatctgattagaanaactcat 4752  
Db 1814 TCAGCGTATAGCTGTGCCAGGTGTACAACTTAAACCAATTCGATTAGAAAACTCAT 1873  
QY 4753 cgagcaatcaatgaacactgcaattatcatatcagagatttcaataacatatatttga 4812  
Db 1874 CGAGCATCAAAATGAAACTGCAATTTATTCATATCAGGATTCATCAATACCATATTTTGA 1933  
QY 4813 aaagcgtttctgttaatgaagagaanaactcagagcagttccatagatgtgcaagat 4872  
Db 1934 AAAGCCGTTTCTGTATATGAAGAGAAACTCACCGAGGAGTTCATGAGATGCAAGAT 1993  
QY 4873 cctgtatcaggtctgcatcgcactcgtccaaactcaataaacttlaattttccct 4932  
Db 1994 CTTGTATCGGTCTGCGATTCCGACTGCTCAACATCAATACCTTATTAATTCCTCT 2053

QY 4933 cgtcaaaaataagttatcaagttagaataacacatagtagtgaactgaatccggttaga 4992  
Db 2054 CGTCAAAAATTAAGTTATCAAGTGAAGAAATCACCATGAGTACAGATCGAGTGAAGA 2113  
QY 4993 atcgcaaaagcttatagtatcttcttcagactgttcaacagccagcattaaagcgt 5052  
Db 2114 ATGGCAAAAGCTTATGATCTTCTTCCAGACTGTGTTCAACAGGCGACGCAATTAAGCTGCT 2173  
QY 5053 catcaaatcactcgtcatcaacaaacggtatcaatcgtgattgtgcgtgagcagagac 5112  
Db 2174 CATCAAAATCACTCGCATCAACCAACCGTTTTCATTCGTGATTTGGCCCTGAGCGAGAC 2233  
QY 5113 gaatacgcgatacgtgtttaaaggacaattacaacaggaatcgaatgcaacggcgca 5172  
Db 2234 GAATATCGGATCGCTTAAAAAGCAATTTCAAAACAGGAATTCAAATGCAACCGGCCCA 2293  
QY 5173 ggaacactgcagcgcatcaacaataattcaaccggaatcaagatattcctcaact 5232  
Db 2294 GGAACACTGCCAGCGATCAACATATTTTCACTTCAATCAGGATTTCTTATTAACCT 2353  
QY 5233 ggaatgctgtttccgggagatcgagtggttagtaaccaatcatcatcagagtagcga 5292  
Db 2354 GGAATGCTGTTTCCGGGGATCGAGTGTGAGTAACCATCATCATCAGAGTACGGA 2413  
QY 5293 taaatgtcttagtgtgcggaagagcataatccgttcagcaggtttagtctgacatct 5352  
Db 2414 TAAATGCTTGATGCTGGAAGAGCATTAATTCCTGACGCCAGTTTACTGACCATCT 2473  
QY 5353 catctgttaacatacttgacaacgctacacttgcacatgttcagaaacaaactgagcat 5412  
Db 2474 CATCTGTATACATCTTGTGCAACGCTACTTGGCATGTTTCAAGAAACACTCGGGCAT 2533  
QY 5413 cgggcttccatacaatcagatagattgtgcacgtatgcccgaacattatcgcgacc 5472  
Db 2534 CGGCTTCCCATFACATGATGATGATGTCACCTGATTTGCCGACATTTATCCGAGGCC 2593  
QY 5473 attatacccatataaatacaagcatcgaatgttgaattaaatcgcggtctcgagcaagc 5532  
Db 2594 ATTTATACCCATTAATAATCAGCATCCATGTTGAATTTATCGCGGCTCGACCAAGAC 2653  
QY 5533 ttcccggttgaatagtgctcaataacacccctgtatcaacggtttagtaagcagagat 5592  
Db 2654 TTTCCGTTGAATATGCTCATTAACACCCCTTGTATTAAGTATTTATTAAGCAGACAGT 2713  
QY 5593 ttatgttcaatgaatataattatctctgtgcaatgtaaacatcagagattttagaca 5652  
Db 2714 TTATTTTCAATGATATATATTTTATCTGTGCAATGTATACATCAGAGATTTTGAACA 2773  
QY 5653 caacgtggtcttcccccccccccaattatgaagcattatcaaggttatgtctcaga 5712  
Db 2774 CAACGTGGCTTTCCCCCCCCCCCCCATTTTGAAGCATTTATCAGGTTATTTGTCTATGA 2833  
QY 5713 ggggatacatattgaattatgaanaataaacaataggggttcgcgcgaacttc 5772  
Db 2834 GCGGATACATATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGGCCACATTTTC 2893  
QY 5773 ccgaaaagtgcacactgacgtcttaagaacaatataatataatgaatcaatcaataaa 5832  
Db 2894 CCGGAAAGTGCACCTGACGCTTAAGAAACCATTAATTAATGACATTAACCTATAAA 2953  
QY 5833 atagcgctatcagagagcccttgcgc 5859  
Db 2954 ATAGGCGTATCACGAGGCGCTTTTCGTC 2980

RESULT 8  
AX138928 5185 bp DNA linear PAT 30-MAY-2001  
LOCUS AX138928  
DEFINITION Sequence 47 from Patent EPI052286.  
ACCESSION AX138928  
VERSION AX138928.1 GI:14274633  
KEYWORDS  
synthetic construct.



[illegible][illegible]



	RESULT	9		
	LOCUS	BDO00693	5185 bp	DNA linear PAT 31-JAN-2002Z
	DEFINITION	Growth hormone and growth hormone-releasing hormone composition.		
	ACCESSION	BDO00693		
	VERSION	BDO00693.1	GI:18623806	
	KEYWORDS	JP 2003350590-A/46. synthetic construct.		
	SOURCE	synthetic construct.		
	ORGANISM	Artificial sequence.		
	REFERENCE	1 (bases 1 to 5185)		
AUTHORS	Morsey,M.A. and George,M.			
TITLE	Growth hormone and growth hormone-releasing hormone composition			
JOURNAL	Patent: JP 2000350590-A 46 19-DEC-2000; Pfizer Prod Inc			
COMMENT	OS Artificial Sequence PN JP 2000350590-A/46 PD 19-DEC-2000 PR 12-Apr-2000 JP 2000111120 PR 12-Apr-1999 US 60/128830 PI MOHAMED ALI MORSEY, MICHAEL GEORGE PC C12N15/09,A01K67/027 ,A6IK31/71L,A6IK38/27 ,A6IK48/00 PC A6IP5/02, PC A6IP43/00,C07K14/60,C12N5/10//((C12N5/10,C12R1:91),C12N15/00, PC A6IK37/36) CC A6IK37/43,C12N5/00,(C12N5/00,C12R1:91)  Key Location/Qualifiers FT source 1..5185 /organism='Artificial Sequence'.			
FEATURES	Location/Qualifiers 1..5185 /db_xref=taxon:32630"			
Source	location/Qualifiers 1..5185 /db_xref=taxon:32630"			

	BASE COUNT	1302 a	1326 c	1260 g	1297 t	
	ORIGIN					
	Query Match	36.7%;	Score 2149.4;	DB 6;	Length 5185;	
	Best Local Similarity	94.5%;	Pred. No. 0;			
	Matches 2293; Conservative	0;	Mismatches	1;	Indels 133;	Gaps
OY	gettcctgcgtacacgagcctcgcttgccgtgcygcttggcgtctgcgcgcgaacggtatcacagt	3625				
Db	554 GCTTCCTCCCTACTACACTCGCTGCGCCTGGTGCCTGCCGTCCGGCGAGCGGTATCACGT	613				
OY	cactcaaaagcggtaatacggttatccacagaatacaggagataaacgcaagaaagacatg	3685				
Db	614 CACTCAAAAGCGGTATFACGGTTATCCAGAATACTAGGGGATTACGACGAAGAAGCATG	673				
OY	tgaagcaaaagccagcaaaaagccaggaaccgtaaagccgcgcttgtgcgcttttc	3745				
Db	674 TCAGCAAAAAGCGCAGCAAAAAGCCAGGAACCCTTAATAAAGCCCGCTTGCTGGCGTTTTTC	733				
OY	cataagctcgcccccttcacagagcatcaaaaaatcgaacgcctcaagtagaggtgagca	3805				
Db	734 CATAGGCTCGCCCCCTTGACGAGCATCACAAAAATCGAGCGCTCAAGTCAGAGCTGGCGA	793				
OY	aaccgcacaggaactataaagataccagcgctttcccccctggaagctccctgtgcgtct	3865				
Db	794 AACCCGACAGACTATTAAGATACCGAGCGCTTCCCCTGGAAGCTCCCTCGGCGCTCT	853				
OY	cctgttcgcacctgcgcgttacccgaataccttgcgccttcccttccttccttcggaaagcgtg	3925				
Db	854 CCTGTTCCGACCCCTGCGCCTTACCGGATACCTTCGCCCTTCTCCCTTCGGGAAGGCTG	913				
OY	gcgccttcataagctcacgcgtglaagtlatcagttcagttcggltagtcgctgcctcaag	3985				
Db	914 GCCTTTCATAGCTCACGCTGTAGGTATCTCATGTTCCGTTGAGGCTGTTCCCTCCAAG	973				
OY	ctggcgctgtgtgcagaaccccccgcttcacgccgaaccgcgcgcgccttatcccgtaaat	4045				
Db	974 CTGGCGCTGTGTGACGAGACCCCGCTTCAGCGAGCCGCTGCGCTTATCCGGTAACAT	1033				
OY	cgtctttagtccaaccccgctaagacaagactlalcgcacattgycgaagcaaccaatgtaac	4105				
Db	1034 CGTCTTAGTCCAACCCGGTAAGACACGACTTATCGCCACATGGCAGCAGCACATGCTGAAC	1093				
OY	aggaatltagcagaagcagaglatgatgagcggtgtctacaaagtctttagaagtgtggtccaaac	4165				
Db	1094 AGGATTTGACGAGCGAGGATGATGTGGCGGTCTTACAAGTCTTGAAGTGGTGCTTAAC	1153				
OY	taggcgtacactagaagaagacagataatttgatactcgcgcctcgtcgaagccagttacctic	4225				
Db	1154 TAGGCGTACACTAANAAGACAGTATTGTGTGATCTGCGCTGTGCGTGAAGCCAGTTACCTTC	1213				
OY	ggaanaaagattgtagtctcttgatccggcaacaacaacacgcgtgtgtagcggtgttc	4285				
Db	1214 GGAAAAAAGATTGGTAGCTCTTGTATCCGCGCAAAACAAACACCGCTGTGAGGGGTGTTTT	1273				
OY	tttgtttcgaagcagcagaaattacgcgcgaagaaanaaagaatccaagaagatcccttgaic	4345				
Db	1274 TTTGTTTCCAAAGCCACAAATTATCGCGCAAAAAAAGATCTCAAGAAGATCCTTTGATC	1333				
OY	ttttcagaaggtctcgaacgctcagtgynaagaaacatcagttaaggatttggatcgt	4405				
Db	1334 TTTTCTACGGGGTGTGACGCTCAGTGAACAAACATCAGTTAAGGATTGTGCTATG	1393				
OY	agatatcataaaaaggaatttcaacctagatcctttt-----	4440				
Db	1394 AGATTATCAAAAAGGATCTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCA	1453				
OY	4441 -----	4440				
Db	1454 ATCTAAAGTATATGAGTAAACTGTGCTGACACTTACAAATGCTTATATCAGTAGGCA	1513				
OY	4441 -----g999g999g999g	4452				



Db	1514	CCCTATCTCAGCGATCTGTCTATTTCGTTATCATAGTTGCCCTGACTCTGGGGGGGGGGG	1573
QY	4453	CGCTGAAGTCCTCGTCTGAAGAAGTGTGTCTGATCATACAGGCGCTGATCGCCCA	4512
Db	1574	CGGTGAGGTGCTGCTGCTGTAAGAAGGTGTGTGCTGACTCATACAGCGCTGAACTGCCCA	1633
QY	4513	TCATCAGCAGAGAAGTGAAGGAGCCAGGCTGTGATGAGGCTGTGTGTAAGTGAAGCAG	4572
Db	1634	TCATATCCAGCCAGAAAGTGAAGGGAGCCAGCTGTGATGAGAGCTTTGTGTTAGTGGACAG	1693
QY	4573	TTGGTGGTTTGTGAACCTTTGTCTTGGCACAGGAACGGTGGCTGTGGGAAGATGGCTG	4632
Db	1694	TTGGTGGTTTGTGAACCTTTGTCTTGGCACAGGAACGGTGGCTGTGGGAAGATGGCTG	1753
QY	4633	ATCTGATCCCTTCAACTCAGCAAAAGTCGATTTATATCAACAAGCGCGCTCCGCTCAAG	4692
Db	1754	ATCTGATCCCTTCAACTCAGCAAAAGTTCGATTTATTCACAAAGCGCGCTCCGCTCAAG	1813
QY	4693	TCAGCGTAAATGCTCTGCGCACTGTTTACAAACCAATTAAACCAATCTCATAGAAAACTCAT	4752
Db	1814	TCACCGTAAATGCTCTGCGCACTGTTTACAAACCAATTAAACCAATCTCATAGAAAACTCAT	1873
QY	4753	CGAGCATCAAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4812
Db	1874	CGAGCATCAAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1933
QY	4813	AAAGCGCTTCTGTGAAGGAGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4872
Db	1934	AAAGCGCTTCTGTGAAGGAGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1993
QY	4873	CCGTGATCGGTCTGCGATCTCGACTCGTCTGCAACATCAATCAACCACTTAACTTCCCT	4932
Db	1994	CCGTGATCGGTCTGCGATCTCGACTCGTCTGCAACATCAATCAACCACTTAACTTCCCT	2053
QY	4933	CGTCAAAAAATTAAGTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4992
Db	2054	CGTCAAAAAATTAAGTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2113
QY	4993	ATGGCAAAAGCTTATGATTTCTTCTTCCAGACTTTGTTCAACAGGCGCATTTACGCTGCT	5052
Db	2114	ATGGCAAAAGCTTATGATTTCTTCTTCCAGACTTTGTTCAACAGGCGCATTTACGCTGCT	2173
QY	5053	CATCAAAATCACTCGCATCAACCAACCGTATTATTCATTCGTGATTTGGCCGTGAGGAGAC	5112
Db	2174	CATCAAAATCACTCGCATCAACCAACCGTATTATTCATTCGTGATTTGGCCGTGAGGAGAC	2233
QY	5113	GAAATAGCGCATTCCTGTTTAAAAAGCAATTTCAAAACAGAAATCAAAATGCAAAACGGCGCA	5172
Db	2234	GAAATAGCGCATTCCTGTTTAAAAAGCAATTTCAAAACAGAAATCAAAATGCAAAACGGCGCA	2293
QY	5173	GGAACACTGCGCAGCGATCAACAATATTTTTCACCTGAAATCAAGATATTTCTTAAATACT	5232
Db	2294	GGAACACTGCGCAGCGATCAACAATATTTTTCACCTGAAATCAAGATATTTCTTAAATACT	2353
QY	5233	GGAAATGCTGTTTCCCGGGAGTGCAGTGTGTGATTAACATCATCAAGAGTACGGA	5292
Db	2354	GGAAATGCTGTTTCCCGGGAGTGCAGTGTGTGATTAACATCATCAAGAGTACGGA	2413
QY	5293	TAAGATGCTGTTGTCGGAAGAGGATCAATTCGTCAGCGAGTTCAGTCGATC	5352
Db	2414	TAAGATGCTGTTGTCGGAAGAGGATCAATTCGTCAGCGAGTTCAGTCGATC	2473
QY	5353	CATCTGTAACATCTTGTGGCAACCTTACCTTTGCAATGTTTCAGAAACAACTCTGGCGAT	5412
Db	2474	CATCTGTAACATCTTGTGGCAACCTTACCTTTGCAATGTTTCAGAAACAACTCTGGCGAT	2533
QY	5413	CGGGCTTCCCATAAATCGATAGATGTGTGCAACCTGATTCGCGGACATTTATTCGGAAGCC	5472
Db	2534	CGGGCTTCCCATAAATCGATAGATGTGTGCAACCTGATTCGCGGACATTTATTCGGAAGCC	2593
QY	5473	ATTATACCATATAAATCAGATCCAGTGTGGAATTAATTCGCGGCTCGAGGAAGCG	5532

[illegible]



Db	794	AACCCGACAGACGATTAAGAGATACAGGCGTTTCCCCCTGGAGAGCTCCCTCGTGGCGTCT	853
QY	3866	ccgtgttcacgacctgcgcgtctacagatactgctccgctctctccctcgggaagcgtg	3925
Db	854	CTGTTCGACACCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTG	913
QY	3926	gcgccttcataagctcaacgcgtgtagtatactaaftctggtggttaagtcgttcgcccaag	3985
Db	914	GGCGTTTCTCATTACTCAGCTCAGGTAGGTATCTCAATTCGATGAGTCTGTTGGTCCAAAG	973
QY	3986	ctgggcctgtgtgacgcgacaccccggtctcaagcccgacgcgcgtgcgcttaccggttaact	4045
Db	974	CTGGGCTGTGTGCACGAAACCCCGCTTACGCCGACCGGTGCGCTTATCCGTTACTAT	1033
QY	4046	cgctcttgatcccaaccgcgttaagaacacgaactatcgccactctggccagcaactgttaac	4105
Db	1034	CGTCTTGGATCCAAACCCGCTAAGCACGACTTATTCGCACTGGCAGCAGCACTGGTAAAC	1093
QY	4106	aggatatacagaagcaggtatgtatgagcgtgtactaagagttcttgaagtgtgacctaac	4165
Db	1094	ACGATTAGAGAGCCAGGATATGTAGCGCGTCTACAGAGTTCTTGAAGTGTGGCCCTAAC	1153
QY	4166	tacgcgtactacagaaagacagatattgtatctcgctctgcgtcgaaagccaattacttc	4225
Db	1154	TACGGCTACACTAGAGAAACAGTATTTGGTATCTGCGCTGTGCAAGCCAGTTACCTTC	1213
QY	4226	ggaataaagtggtgtgagctgtacgtcttgatccgcgcaacaacaacacgcgttgtagcgtgtt	4285
Db	1214	GGAAAAAGAGTTGGTAGCTCTTGATGCCGCAACAAACACCGCTGAGGCGGTGTTT	1273
QY	4286	tctgtctgaacagcagcagatatacgcgcgcagaaaaaagatctcaagaagaatcccttcatc	4345
Db	1274	TTTGTTCAGACGACGACAGATTACGGCGCAAAAAAGATCTCAGAAAGATCCTTTGATC	1333
QY	4346	ttttccacgggtgtcagcgtctagcgtgagaaacgaaactacgcttaaggagatttggtaatg	4405
Db	1334	TTTTCTACGGGGTCAAGCTCAGTGGAGACGAAACTCAGTTAAGGATTTTGGTCAATG	1393
QY	4406	agatatacaaaaagatctctacatagatccctt-	4440
Db	1394	AGATTATCAAAAAAGATCTTACCTGTATGTCCTTTAAATTAATAAATGAAGTTTAAATCA	1453
QY	4441	-	4440
Db	1514	CCTATCTCAGGATCTGTCTATTTGTTTCATCATAGTTGCTGACTCGGGGGGGGGGG	1573
QY	4453	cgctgaggtcgtgcgtctgtagaagaagtggtctgcgaactacatacagccttgaatcgccca	4512
Db	1574	CGCTAGGTCGTCCCTCGTAGAAGAGTGTGTGACTCATACAGGCGCTGAATCGCCCA	1633
QY	4513	tcaatcacgcagaagltgagggagccaagtttagagagccttgttgttagtggtggccag	4572
Db	1634	TCATCCAGCAGAAAGTAGGAGACCCAGGTGATGAGACCTTGTGTAGTGAGACAG	1693
QY	4573	ttgtgtatttgaacttttgttttgccagcggaaacgcgtgcgtctgcggaaagatgcgtg	4632
Db	1694	TTGGTGATTTTGAACCTTTTGGCTTTGCCACGAAAGGCTGCGTGTGGGAAGATGGTG	1753
QY	4633	atctataccttcaactcagcaaaagtttgcatttatacaaaagcgcgcgtccgcgttaag	4692
Db	1754	ATCTGATCCTTCACTCAGCAAAAGTTGATTTATTCACAAACCGCGCGTCCGTAAAG	1813
QY	4693	tcaagcgtlaatgcctgcgcaggtttatacaaccaatlaaccaattctgatttagaaaaactcat	4752
Db	1814	TCAGCGTAAATGCTGTGCCAGTGTTCACACCAATTAACAAATTCTGATTTAGAAAACTCAT	1873
QY	4753	cgagcgtcaaatgaacacgtcaattattatlaatacagattalcaatatcaataatttggaa	4812
Db	1874	CGAGCATCAATTAACACTGCATTTTATTTATATTCAGAGATTAATCAATACCATATTTTGA	1933

QY	4813	aaagccgtcttcgtatagtgaagggaaaactccagggcagcttccatgagat	4872
Db	1934	AAACCCCTTCTGTAAAGAGGAAAAACTCACCCAGGCACTTCCATGAGATGGCAAGAT	1993
QY	4873	ccctgctatcgctcgcgactccgactcgctccaacatcaatatacctaattatccct	4932
Db	1994	CTGTGATCGGTCTCGGATTCGCACTCGTCCAACTCAATCAATAACCTATTAAATTCCTCT	2053
QY	4933	cgctaaaataaggtttctcaagtgagaataccactgaatgaagactgaatccggtlaga	4992
Db	2054	CGTCAAAATTAAGGTATTCAAGTAGAAATCACCATGATGAGAGACTGAATCCGGTAGA	2113
QY	4993	atggaanaagcttatgcatcttcttccagactgtgtcaacaggccagccattacgctgt	5052
Db	2114	ATGGCAAAAGCTTATGCACTTTCTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTGCT	2173
QY	5053	catcaaatcaactcgcatcaacaacacgattatcattcctgtgtatltcgccgtgagagac	5112
Db	2174	CATCAAAATCATCGCATCAACCAACCCGTTATTCTCGATGTGGCCTGAGCAGAGAC	2233
QY	5113	gaatacgcgactcgctgtctaaaggaataatacaacagaagatcgaatgaaacccgcga	5172
Db	2234	GAAATTCGGATCGCTGTTTAAAGAGCAATTATTCAAACAGGATGGAATGGAACCGCGCA	2293
QY	5173	ggaacactcgccggcgatcaacaacatattcaactgtaacgaatatcttctaact	5232
Db	2294	GGAAACCTGCGAGCATCAACAAATTTTACCGTGAATCGAGATTCCTGTATATACCT	2353
QY	5233	ggaatgctgttctccggggatcgcaatgtgtgagtgaacctgacatcaatgaagaa	5292
Db	2354	GGATGCTGTTTTCCGGGGATTCGAGTGTGTAGTATACCATGCACTATCAGAGATTACGGA	2413
QY	5293	taaaatgctgtagtgcggaagaagagataaattccgtacgacgaagtaagttagttgacatc	5352
Db	2414	TAAATGCTTGAATGTCGGAAGAGGCAATTAATTCGTGAGCCAGTTAGTGTGACATCT	2473
QY	5353	catctgtaacatcattggcaaacgctaccccttgcactgttcaatgttgaanaaacctctgcgat	5412
Db	2474	CATCTGTATACATCATTTGGCAACCTCACTTTGGCATGTTTCAGAAACAACCTGGCGCAT	2533
QY	5413	cgggctcccatcaatcgaatagatgtgcgacctgatttgcgcgacattatcggaagcc	5472
Db	2534	CGGGCTTCCATPACATCGATAGATGTGCGACCTGATTTGCCGACATTATGCGAGGCC	2593
QY	5473	attatacccatataaataagacatcaatgttgaattcaatcgcgcgcctcgagaacag	5532
Db	2594	ATTATATACCATTAATAATCAACATCCATGTTGGAATTTATGCGCGCTCGAGCAAGCG	2653
QY	5533	ttcccgcttgtaataatggtcaaacacccctgtattactggttaagtgaagaacagt	5592
Db	2654	TTTCCCGTGTGAATATGCTCTATACACCCCTGTATTACTGTTATGTATAGCAGACAGTT	2713
QY	5593	ttatgttcatagatagatatatttcaactgtgtgaatgttacaatcaagaattttagaca	5652
Db	2714	TTATTTGTATGATGATCATATATTTTATCTGTTGCAATGTAACATCGAGATTTTGGAGA	2773
QY	5653	caacgttggtcttcccccccccccaattatgaagacatttaaggggttatgtctcatga	5712
Db	2774	CAACGGGTGCTTCCCCCCCCCCCAATTATTAACCATTTATACAGGTTATTGTCTATGA	2833
QY	5713	ggcgatgaactattgaatgattagaanaataacaacaataggggttcgcgcacattc	5772
Db	2834	GCGGATACATATTGATATCTATTTGAAAAATAACAATATGGGGTTCCCGCACATATTC	2893
QY	5773	cccgaaaatgtgcacactgaagctctaagaacaacattatatcatatgaacttaaccataaa	5832
Db	2894	CCCGAAAGGCGCACCTGACGCTTAAGAAACCATTTATATCATATACATTAAACCTATAAA	2953
QY	5833	atagggtatcaagagcccttctgc	5859
Db	2954	ATAGCGTATACAGAGCCCTTTCGC	2980



[illegible]

Dd	974	CTGGGCTGTGTGACGAAACCCCGTTTCAGCCCGACCCGTGGCGCTTATCCGGTAACAT	1033
Qy	4046	cgctctgagtcaccaaccgcytgaagacagactatcgccactgycagcagcactgyltaac	4105
Dd	1034	CGTCTTAGTGTCAACCCGGTAAACACAGACTTATCGCCACTGGCAGCACACTGTGTAC	1093
Qy	4106	aggattcgagagcgaggtgtgtagcggtgtcctaagagttctttaaagtgltgycctaac	4165
Dd	1094	AGGATTAGCAGAGGAGGATGTATGAGCGGGCTACAGAGTTCTTAAAGTGGGGCTAAC	1153
Qy	4166	tacggtacacactagaagaagaaacttggltatctgycctctgtgaagccaglttaacctc	4225
Dd	1154	TACGGTACACTAGAAAGACAGTATTGGATCTGGCTCTCTGAAGCAGATTACCTTC	1213
Qy	4226	ggaaaaagatltgtagctctgtatccgcgcaaacacacacccgcgtgtagcgtgttctt	4285
Dd	1214	GGAAAAAGATGTGTACTCTTGTATCCGGAAACAAACACCGCTGTGAGCGGTGGTTTT	1273
Qy	4286	tttgttgcagcagcagacttacgcgcagaaaaaagatctccaaagaaagatccctgtgc	4345
Dd	1274	TTTTTTTSCAAGCAGCAGATTACGCCAGAAAAAAGATCTCAGAAAGATCTTTATATC	1333
Qy	4346	tttctcacgggtgtcgcgtcscgtggaagaaactcaglttaaggatttggltcatg	4405
Dd	1334	TTTTCTACGGGGTCTGCAGCTCAGTGGAAAGAAACTCAGTTAAGGATTTTGGCTATG	1393
Qy	4406	agattatcaaaaaagactcttcacctagaatccctt-----	4440
Dd	1394	AGATTATCAAAAAAGATCTTACCTAGATCTTTTAAATTTAAAAATGAAGTTTAAATCA	1453
Qy	4441	-----	4440
Dd	1454	ATCTAAAGTATATATGATGAATAACGTGGTGTGACAGTTACCAATGCTTAATCAGTAGCA	1513
Qy	4441	-----ggggggggggggg	4452
Dd	1514	CCTATCTCAGCAGATCTGTCTATTTCGTTCATCCATAGTTGCTGTACCTGGGGGGGGGG	1573
Qy	4453	cgctggaagttctgcctctgaaagaagtggtctgcactaatacagagccgtaactgcacca	4512
Dd	1574	CGGTAGGTGTGGCTCGTGAAGAAGGTGTGCTGACTATACACAGCGCTGAATGCCCCA	1633
Qy	4513	tcaatccagccaagaatgtaaggagaccaaggtgtgataagagcttgttgaagtgcagac	4572
Dd	1634	TCATCCACGCCGAAAGTAGAGGAGCCACGGTTGTATAGAGCTTTGTGTAGGTGACACAG	1693
Qy	4573	ttgtgatttgaacttgcctcttgcacagaaacggtctgcgcttctcggaagaatgagtg	4632
Dd	1694	TTGGTGAATTTGAACCTTTTGCTTGGCACGGAAGCGTCTGGTGTGCGGAAGATGCGTG	1733
Qy	4633	atctgatccctcaactcagacaaaagtctgatttatcaacaaagccgcgcgtccgtcaag	4692
Dd	1734	ATCTGATCCTTCACTCAGCAAAAAGTTGATTTATCAACAAAGCGCGCTCCCTCAAG	1813
Qy	4693	tcaogctaaatgctctgcgcagtggttaaaccaaatlaaccgaattctgattagaaaaaactat	4752
Dd	1814	TCAGGTAAATCTCTGTGCAGTGTACACCAATTAACCAATTTCTGATTAGAAAACTCAT	1873
Qy	4753	cgaagatcaaaatgaactgcgaatttatataatcagagttatacaatatcataatttctgaa	4812
Dd	1874	CGAGATCAAAATGAAACTGCCATTATTATCTATACAGAGATTATTCANATACATATTTTGA	1933
Qy	4813	aaagccgttctgtlaatgaagaggaanaactcacagagcgagttcgaatgaatgycgaagat	4872
Dd	1934	AAAGCGCTTTCTGTAAAGAGGAGAAACACACCGAGCGAGTTCATATGAGATGCGCAGAT	1993
Qy	4873	ccgtgtatccggtctgycgattccgcactgcgtccaacatcaataaacactttaaatttccct	4932
Dd	1994	CCTGGTATCCGGTCTTCGCATTCCGACCTGTCTCAACATCATATACACCTATTAATTTCCCT	2033
Qy	4933	cgtcaaaaataaagttatacaagttagaanaatcaccaatgagtgacgactgaatccggtgaga	4992
Dd	2034	CGTCAAAATTAAGTGTTCCTCAAGTAGAAATACCATGATGACACTGAATCCGCTGAGA	2113



[illegible][illegible]











Db	2654	TTTTCCCGTTGGAATATGGCTCATACACCCCTTGTAATTAAGTGTATATGTAAGCAGACAGTT	2713
QY	5593	ttatctgtcatgaaatgataatacttcatctctgtgcaatgtaacataagagatttgaaga	5652
Db	2714	TTATTTGTTCAAGATGATATATTTTATCTGTGTGCAATGAATCAGAGATTGTAGACA	2773
QY	5653	caacgttgagcttcccccccccccaatttgaagcatttatacgaagtattgtctcatga	5712
Db	2774	CAAGGTGGCTTTCCCCCCCCCCCTTATTGAAACATTATTACGGGTTATTGTCTCATGA	2833
QY	5713	ggcgatacatatttgaatgattttagaaaaataacaatatggggttcogcgacatttc	5772
Db	2834	GCGGATACATATTCTTAAGTATTATTAGAAAAATAAACAAATATAGGGTTCCGCCACATTTTC	2893
QY	5773	cccgaaagctgcacacttgcactctagaagaaccattatctatctagaattaaaccttaaa	5832
Db	2894	CCCGAAAGTGCCACACTTACGCTAAGAAACCATTTATTATCTATGACATTAACTTAATAA	2953
QY	5833	ataagcgatatacaagaagcccttcgctc	5855
Db	2954	ATAGCGCTATCACGAGGCCCTTTCCGTC	2980

RESULT	14				
LOCUS	AX080955				
DEFINITION	AX080955	4779 bp	DNA	linear	PAT 27-FEB-2001
ACCESSION	AX080955	Sequence 5 from Patent WO0109303.			
VERSION	AX080955.1	GI:13169892			
KEYWORDS					
SOURCE					
ORGANISM		synthetic construct.			
REFERENCE		synthetic construct			
AUTHORS		artificial sequence.			
TITLE		1 (bases 1 to 4779)			
JOURNAL		Hermanson, G. G.			
FEATURES		Flt-3 ligand-encoding polynucleotide as a polynucleotide-based			
source		vaccine enhancer			
		Patent: WO 0109303-A 5 08-FEB-2001;			
		VICAL INCORPORATED (US)			
		Location/Qualifiers			
		1..4779			
		/organism="synthetic construct"			
		/db_xref="taxon:32630"			
		/note="cDNA"			
BASE COUNT	1219 a 1203 c 1139 g 1218 t				
ORIGIN					

Query Match	36.6%	Score 2146.8	DB 6	Length 4779
Best Local Similarity	94.4%	Pred. No. 0		
Matches 2292	Conservative	0	Mismatches 2	Indels 134
				Gaps 1
OY 3566	gcttcctcgctaacgactcgcgtcgtcgtcgtcgcgttcgcgtcgtcgcgcgcgcgtatcaagct	3625		
Db 2108	gcttccctcgctcactacgacacgctgcgcgtgcgcgtgcgttcgcgttcgcgcgcgcgcgcgcgcgt	2167		
OY 3626	cactcaaaagcggtatcgcgtcgtatccacaaagataaggagataaagcagaagaagaacatg	3685		
Db 2168	cactcaaaagcgcgatnattggtttatccacgaatcaggggatatacccaagaaagaacatg	2227		
OY 3686	tgaagcaaaagcgacgaacaaagccaggaacgtaaaagccgcgcttctgtcgttttc	3745		
Db 2228	tgagcaaaaagccacacaaaagccacaggaaccgtaaaaaagccgcgcttcgtgcgcttttgc	2287		
OY 3746	catgaagcccgcccccttgaagagatcatcaaaaatcgaacgcttcaagttaaggtatggca	3805		
Db 2288	catagagcttcgcccccttcacagacatccaaaaatccagccgtcaanagtacaggttgcca	2347		
OY 3806	aaccgacagagactataaagataccaagcgctttccctctgaagatccctctgtcgtct	3865		
Db 2348	aaccgacagagactttaaagataccagagcggtttccctctggaagctccctctgcgcctct	2407		

OY	3866	ccggtccgaaccctgcgcgcttacccgaatacctgtccgcctctctcccttcgggaagcgt	39226
Db	2408	cctgttcgcgaccctgcgcgcttacccgaatacctgtccgcctctctcccttcgggaagcgtg	24676
OY	3926	ggccttctcaataagctaacgcctgttaagtaactcaatcttcggtgttaagtctgtccgaag	39856
Db	2468	ggccttctctcaatgcctcaacgctgttaagtaactcaatcttcggtgttaagtctgtccgaag	25226
OY	3986	ctgggcctgtgtgcacgaaccgccggttccaaagcccgaccgcctgtgcgctctcaacgtaatac	40456
Db	2528	ctgggcctgtgtgcacgaaccgccggttccaaagcccgaccgcctgtgtatccggttaacat	25876
OY	4046	cgctctgaagctcaaaccccggttaagaacacgactatctgcacatgtgcagcaacatgttaac	41056
Db	2588	cgctctgaagctcaaaccccggttaagaacacgactatctgcacacactgtgcacacactgttaac	26476
OY	4106	aggaattgcagagcgaagcgtgtgttagcgggtgtcctcaaggaagttcttgaagtggtgcctaac	41656
Db	2648	aggattgtacacagcgaagcgtgtgttagcgggtgtcctcaaggaagttcttgaagtggtgcctaac	27076
OY	4166	tacgcctcaacataagaagaacgaattctgtgtatctgtgcctgtcgtgaagcgaataccttc	42226
Db	2708	tacgcctctacactataagaagacgaatttgggtatctggcgtctcgtgaagcgaataccttc	27676
OY	4226	ggaaaaaagatgtgttagcctcttgatccgcgcaaaacaaacacgcctgtgttagcgggtgttc	42856
Db	2768	ggaaaaaagatgtgtgttagcctcttgatccgcgcaaaacaaacacgcctgtgttagcgggtgttc	28276
OY	4286	tttgtttgcaagcgcgaattatcgcgcgaaaaaaaggatctcaagaatactccttgttc	43436
Db	2828	tttgtttgcaagcgcgaattatcgcgcgaaaaaaaggatctcaagaatactccttgtgtac	28876
OY	4346	ttttctacggggtcttgacgcctcgaatgtgaacgaanaacatcgcgttlaaggaattctgtcatg	44056
Db	2888	ttttctacggggtcttgacgcctcgaacgctcaatggaacgaanaacatcgaatttttggtcatg	29446
OY	4406	agatatacaaaaagatcttcaactagatccctt-----	44406
Db	2948	agattatcaaaaagatcttcaactagatccctttaaatttaaaatgaagttaaataca	30076
OY	4441	-----	44406
Db	3008	atcttaaaatattatgatgaataaacttggcttgacagattaccatgcttaattacagtgaagca	30676
OY	4441	-----	44576
Db	3068	ccctatctcagcgaatctgtcttaatttcgttccatccatgaatttgcctgactccggggggggggg	31226
OY	4452	ggcgtcgaagctcgtgcctgtgaagaaggctgtctgtactcaatacagccttgatccgccc	45116
Db	3128	ggcgtcgaagctcgtgcctgtgaagaaggctgttgcctgaactcatatcacgcttgaaatggcccc	31876
OY	4512	atcaatcagacgaagaatgtgaaggacgaagctgtgatagaagacttgtgttlaagtggaacca	45776
Db	3188	atcaatccagcgaagaatgtgaaggacgaagcctgtgatagagacttgtttagtgatgacaca	32476
OY	4572	gttgttgaatttgaaacttttgcttgccttgcaaggaagcgtctgcgttgttcgggaagatgcgt	46336
Db	3248	gttgtgtgaattttgaaacttttttgccttgccttgcaaggaagcgtctgcgttgttcgggaagatgcgt	33076
OY	4632	gattcgaacccctcaactcaagaagaatctgattatattcaacaagcgcgcgcctccgtcaaa	46936
Db	3308	gattcgaatccttcaaaactcagcaaaagtctgattattatcaaaagcgcgcgcctccgtcaaa	33676
OY	4692	gtcagcgttaatgtcctcgcgaagtgttacaaccaatlaaccaatctcgtatagaanaaaactca	47556
Db	3368	gtcagcgttaatgtcctcgcgaagtgttacaaccaatlaaccaatctcgtatagaanaaaactca	34226
OY	4752	tgcagcctcaaatgaanaatgtgaatttcaatacgaagatatacaataacataattttga	48116
Db	3428	tgcagcctcaaatgaanaatgtgaatttattatattacaggaattatcaatattcatttttttga	34876
OY	4812	aaaaacgctttctcgtlaatgaagggaanaactcacgcgagcgaagtctcataggaatgtgcaga	48716







```

Db 3097 GGAATAAGTTGAGTCTGTGATCCGGCAAAACACCGCTGGTGGTATT 3156
Qy 4286 ttgtttcgaagcaagatcttcgacgcaagaaaaagatctccaagaagatccctttgac 4345
Db 3157 TTTGTTTCCAAACACACACATTAAGCGCAGAAAAAAGATCCAAAGAAATCTTTTGATTC 3216
Qy 4346 ttctcaggggctcgaagcagtcagtgaaacgaactcagcttaagggatcttctgcatg 4405
Db 3217 TTTTTCAGGGGCTCTGACGCTCAGTGGAAAGAAACTCAGTTAAGGATTTTGGTCATG 3276
Qy 4406 agattatcaaaaaaggatcttcacctagatcctt----- 4440
Db 3277 AGATTATCAAAAAAGATCTTCACTAGATCCTTTAAATTAATAAAGAGTTTAAATCA 3336
Qy 4441 ----- 4440
Db 3337 ATCTAAGATATATAGTAATAAATTGTCCTGACACTTACCAATGCTTAATCAAGTAGCA 3396
Qy 4441 ----- 4451
Db 3397 CCTATCTCAGCATCTGCTTATTGTTTCATCCATAGTTGCCGACTCCGGGGGGGGGG 3456
Qy 4452 ggcgtgaagctcgtcgtcgtgaagaaagtgctgactcaataccagccctgaatcgcgcc 4511
Db 3457 GCGCTGAGGCTGCTCGTGAAGAAAGTGCTGCTGACTCATACGAGGCTGAAATCGCCCC 3516
Qy 4512 atcatccagccgaagaaagaaagaaacaggttgatgaagcttggctgtaagtgagca 4571
Db 3517 ATCATATCCAGCCGAAAGTGAAGGAGCCAGGTTGATGAGAGCTTTGTTAGCTGAGCA 3576
Qy 4572 gtctgtgatttgaactttgtcttgcagagaaagctgctgctgtctgaggaagatgct 4631
Db 3577 GTTGGTGATTTGAATCTTTGCTTGTGCCAGGAAAGCGTGTGCGGAAGATGCGCT 3636
Qy 4632 gatctgactcctcaactcagcaaaagttcgatttattcaacaagccgcgtccgctcaa 4691
Db 3637 GATCTGATCTCTCAACTCAGCAAAAGTTGATTTATTCAAAAGCCGCCGCTCCGCAA 3696
Qy 4692 gtcacgttaatgtctgcgaaggttacaacaaacttaaccatcttgatagaanaactca 4751
Db 3697 GTACACGTATGCTCTGCGAGGTGTAACCAATTAACCAATTCGATTGAGAAAACTCA 3756
Qy 4752 tcgagcatcaaaatgaactgaactatcatcatcagaatcaataccatatttga 4811
Db 3757 TCGACATCAAAATGAACGCAATTTATCATATCGATGATTCATATTTTGA 3816
Qy 4812 aaaaagcgttctctgaatgaagaaagaaactcaacgaagcagtlccatagatgagca 4871
Db 3817 AAAAGCCGTTTCTGTAATGAAGAGAAACTCACGAGGCAAGTTCCATGAGATGCAAGA 3876
Qy 4872 tcttggtatcgttctgagatcttcgaactgttcaacatcaatacaacttaattccgc 4931
Db 3877 TCTGTGATCGGTCTGCGATTCGACTCGTCCAACTACATACACTTAATTTTCCCC 3936
Qy 4932 tcgtcaaaaaatgaagttatcaagtgagaatcacatcagtgagtgagactgaatccgctgag 4991
Db 3937 TCGTCAAAAAATGAAGTTATCAAGTGAAGAAATCACCATGAGTGAATCCGGGTGAG 3996
Qy 4992 aatggcaaaagcttatgcatcttcttccagactgttcaacagggccagcattcagctcg 5051
Db 3997 AATGGCAAAAGCTTATGCAATTTCTTCCAGACTGTTCAACAGGCCACGCAATTAACGCTCG 4056
Qy 5052 tcataaaatcaactcgaatcaaaccaacgcttatcatctgtagtgcgctgaagcgaga 5111
Db 4057 TCATCAAAATCACTCGCATCAACCAACCGTTATTCATTCTGATGATCGCTGAGCGAGA 4116
Qy 5112 cgaatatcgcgactcgtctttaaagagaaatatacaaacagaaatcgatgcaacccgagc 5171
Db 4117 CGAAATAGCGCATCGCTGTTAAAGAGCAATTTACAAACGAAATGCAATGCAACCGGCG 4176
Qy 5172 aggaacactgcagcgcatcaacaatatcttcaactgaatcaagatatcttcttaaac 5231

```

```

Db 4177 AGGAACATGCTCCAGCGCATCAACAATATTTTCACTGATCAAGATATCTTCAATACC 4236
Qy 5232 tggatgctgttctccgggagtcgaatggtgagtaacatgcatcatcagaagaaagc 5291
Db 4237 TGGATGCTGTTTTCCCGGGGATCGCATGTGTAGTAACCATCAATCAGAGTAACGG 4296
Qy 5292 ataaaaatcgtatgctcgaagaaagcaataatctcgtcagccagctttagtcgaacatc 5351
Db 4297 ATAAATGCTTGAATGTCGGAAGAGGCATTAATTCGTCAGCCAGTTTGTGTGACATC 4356
Qy 5352 tcaatcgtacaatcatctgcaacgctacacttgcaatgtttcagaacaaacttgcgca 5411
Db 4357 TCATCTGTAAATCATTTGGCAACGCTACCTTGCATGTTTCAGAAACAACCTTGCGCGA 4416
Qy 5412 tcgggcttccatcaaatcgaatagatgttcgaacccgattggcccgacattatcgagacc 5471
Db 4417 TCGGCTTCCCATACAAATGATGATTTGTGCACTGATTTGCCGACATTTATGCGAGCC 4476
Qy 5472 catltaacatataatcaatcagcatcgaatgttgaaatlaatcgcgccctcgaagcaagc 5531
Db 4477 CATTTATACCATATAATCAGCATTCATGTTGAATTTAATCGCGGCTCGAGCAAGAC 4536
Qy 5532 gtttccgcttgaataatgctcacaacacccctgtatctactgtttatgtaagcaagc 5591
Db 4537 GTTTCCTGTTGAATATGCTCATTAACACCCCTGTATCTGTTATGTTAAGCAGACAGT 4596
Qy 5592 ttatgttcataatgatatatttatacttctgcaatgaaatcaagaatttgaagc 5651
Db 4597 TTTATGTTCAATGATGATATATTTTATCTTGTGCANGTGAATCAAGAAATTTGAGAC 4656
Qy 5652 acaaagtgcttccccccecccaattatgaagcatlatacagggttatgtctcatg 5711
Db 4657 ACAAGTGGCTTTCGCCCCGCCCATTTATGAAGCATTTATCAGGGTTATGTCTCATG 4716
Qy 5712 agcgaatacatctgaatgtatctagaanaataaacaataggggttccgcgacatt 5771
Db 4717 AGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCGCCACATTT 4776
Qy 5772 ccccgaaaagtgccactgaagcttaagaaacatattatcaatgaatcaacttaaa 5831
Db 4777 CCGGAAAGTGCCACCTGACGCTTAAGAAACCATTTATTTATCATGACATTAACCTTAA 4836
Qy 5832 aatagcglatcacgagagcccttctgctc 5859
Db 4837 AATAGCGTATCAAGAGGCCCTTTGCTC 4864

```

Search completed: August 20, 2002, 22:21:33  
 Job time: 15831 sec



THIS PAGE BLANK (USPTO)



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 19:03:32 ; Search time 563.56 Seconds  
(without alignments)  
17849.744 Million cell updates/sec

Title: US-09-826-206-3  
Perfect score: 5859  
Sequence: 1 tcgcgcgttcggtgatgac.....taccagcgcccttcgtc 5859

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_032802:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5859	100.0	5859	24	AAS18308	DNA sequence of sh
2	2171.2	37.1	3604	22	AAH79219	Human leukocyte al
3	2149.4	36.7	4867	22	AAD09428	Vaccine vector V1J
4	2149.4	36.7	4867	22	AAD09614	Vaccine vector V1J
5	2149.4	36.7	4945	22	AAD09421	Vlins-tPA vector t
6	2149.4	36.7	4945	22	AAD09607	Vlins-tPA vector t
7	2149.4	36.7	5108	22	AAC86258	Plasmid GHRH1-29WT
8	2149.4	36.7	5108	22	AAC86259	Plasmid GHRH1-29WT
9	2149.4	36.7	5111	22	AAC86262	CMV IE promoter eh

10	2149.4	36.7	5185	22	AAC86254	pGHRH-4 construct.
11	2149.4	36.7	5188	22	AAC86266	Plasmid GHRH1-44W
12	2149.4	36.7	5254	22	AAC86267	Plasmid pGHRH1-44W
13	2146.8	36.6	4779	22	AAF30340	Expression plasmid
14	2146.8	36.6	4862	22	AAS09198	DNA sequence for e
15	2146.8	36.6	4864	16	AAV01114	Expression vector
16	2146.8	36.6	4864	19	AAV21737	Vlinoe sequence us
17	2146.8	36.6	4864	20	AAV57727	Human VEGF(145) co
18	2146.8	36.6	4864	22	AAD09427	Vaccine vector V1J
19	2146.8	36.6	4864	22	AAD09437	Vaccine vector V1R
20	2146.8	36.6	4864	22	AAD09613	Vaccine vector V1J
21	2146.8	36.6	4864	22	AAD09623	Vaccine vector V1R
22	2146.8	36.6	5259	20	AAH81561	Cytokine-encoding
23	2146.8	36.6	5259	20	AAH81565	Cytokine-encoding
24	2146.8	36.6	5310	22	AAF30314	Human Flt-3 ligand
25	2146.8	36.6	5322	20	AAH81563	Cytokine-encoding
26	2146.8	36.6	5322	21	AAH63225	Plasmid VR4151, co
27	2146.8	36.6	5322	22	AAH30313	Mouse Flt-3 ligand
28	2146.8	36.6	5422	20	AAH81564	Cytokine-encoding
29	2146.8	36.6	5428	20	AAH81560	Cytokine-encoding
30	2146.8	36.6	5469	20	AAH81562	Cytokine-encoding
31	2146.8	36.6	5480	20	AAH81577	Cytokine-encoding
32	2146.8	36.6	6709	22	AAF30339	Bicistronic plasmid
33	2146.8	36.6	6729	22	AAF30341	Bicistronic idioy
34	2146.8	36.6	7099	22	AAH22814	DNA sequence of a
35	2146.8	36.6	7285	20	AAH89797	DNA of pVR 1012-GP
36	2146.8	36.6	7285	20	AAH59392	Plasmid pVR 1012-GP
37	2146.8	36.6	7528	22	AAF30316	Bicistronic idioy
38	2146.8	36.6	8366	22	AAH22810	DNA sequence of a
39	2145.2	36.6	7073	20	AAH59391	Plasmid pVR 1012-G
40	2145.2	36.6	7272	20	AAH89798	DNA of pVR 1012-sg
41	2145.2	36.6	7521	22	AAF30315	Bicistronic chimera
42	2143.6	36.6	4864	15	AAO74207	Expression vector
43	2143.6	36.6	7073	20	AAH89796	DNA of pVR 1012-GP
44	2143.6	36.6	7272	20	AAH59393	Plasmid pVR 1012-s
45	2143	36.6	4912	21	AAV71402	Plasmid VR-1012 DN

ALIGNMENTS

RESULT 1	
ID AAS18308	standard; DNA: 5859 BP.
AC AAS18308:	
XX	
DT 12-MAR-2002	(first entry)
XX	
DE DNA sequence of shuttle vector plasmid pADM291-4.	
XX	
KW Cloning vector: Ketogulonigenium replicon: endogenous plasmid: mutant;	
KW transformed host cell: Escherichia coli; pADM291-4; circular; cyclic;	
KW shuttle vector plasmid; ds.	
XX	
OS Ketogulonigenium sp. strain ADM291-19.	
OS Synthetic.	
XX	
PN WO200177347-A2.	
XX	
PD 18-OCT-2001.	
XX	
PF 05-APR-2001; 2001WO-US11059.	
XX	
PR 05-APR-2000; 2000US-194625P.	
XX	
PA (ARCH ) ARCHER-DANIELS MIDLAND CO.	
PA (DELI/) D'ELIA J.	
XX	
PI D'Elia J;	
XX	
DR WPI; 2002-049150/06.	
XX	



PT Novel nucleic acid vector comprising Ketogulonigenium replicon found on  
PT a specific deposited endogenous plasmid, useful for producing  
PT polypeptides and/or transcripts by culturing host cells transformed  
PT with vector -  
PS Claim 3; Fig 3; 66pp; English.  
XX  
XX  
XX The present invention relates to the isolation of vectors comprising  
CC a Ketogulonigenium replicon found on the endogenous plasmid, PADM291.  
CC The invention also describes methods of transforming host cells with  
CC the vectors and producing polypeptides and/or antisense transcripts by  
CC culturing the transformed host cells. The vectors are useful for  
CC transforming a host cell by conjugation or electroporation.  
CC The vectors which have a replicon functional in both Ketogulonigenium  
CC and Escherichia coli, enable the cloning of certain genes of  
CC Ketogulonigenium in E.coli as the latter is an efficient host for  
CC amplification of vector DNA. The present DNA sequence represents the  
CC shuttle vector plasmid PADM291-4.  
XX  
XX Sequence 5859 BP; 1444 A; 1487 C; 1566 G; 1362 T; 0 Other;  
SQ  
Query Match 100.0%; Score 5859; DB 24; Length 5859;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5859; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 tcgcgcgttcctcgatgaacggtgaacacccctgacacatgcagctcccgagagctga 60  
DB 1 tcgcgcgttcctcgatgaacggtgaacacccctgacacatgcagctcccgagagctga 60  
QY 61 caactgtctgtaagcgatgcgcgagacagacaccccgtaagcgcgctgaacggtg 120  
DB 61 caactgtctgtaagcgatgcgcgagacagacaccccgtaagcgcgctgaacggtg 120  
QY 121 ttgcgcggtgctcgagcgctggtcttaactatagcgcatcagacagatgttactgagagtg 180  
DB 121 ttgcgcggtgctcgagcgctggtcttaactatagcgcatcagacagatgttactgagagtg 180  
QY 181 accatagcggtgtgaataaccgacagatcgtaagagagaataaccgatcagcgcc 240  
DB 181 accatagcggtgtgaataaccgacagatcgtaagagagaataaccgatcagcgcc 240  
QY 241 attcgacatcagcgctgacgaactgttggaagggcgatcggtgcgggctcttcgcat 300  
DB 241 attcgacatcagcgctgacgaactgttggaagggcgatcggtgcgggctcttcgcat 300  
QY 301 taagcgcagctgagcaaaagggggtgctgctcaagcgatgaatttggtgaacgcaggt 360  
DB 301 taagcgcagctgagcaaaagggggtgctgctcaagcgatgaatttggtgaacgcaggt 360  
QY 361 ttcccgacatcagcgctgtgaataacgacgagcagtgaaattcggaattca 420  
DB 361 ttcccgacatcagcgctgtgaataacgacgagcagtgaaattcggaattca 420  
QY 421 tagaatttgtgtagagtgctgacgctctgacagaggggtgctgcggagagatctcgtgt 480  
DB 421 tagaatttgtgtagagtgctgacgctctgacagaggggtgctgcggagagatctcgtgt 480  
QY 481 ctcaagtagggcgacaatagagaggtgttagttgccccctatcgcgtctcgtcgtagcg 540  
DB 481 ctcaagtagggcgacaatagagaggtgttagttgccccctatcgcgtctcgtcgtagcg 540  
QY 541 catlgtgcatcctgcccgcgacatagtatccgctagagagatcagatagattcttcgc 600  
DB 541 catlgtgcatcctgcccgcgacatagtatccgctagagagatcagatagattcttcgc 600  
QY 601 ctgctcgaggtgctgaggtctgctgcgagctgctgcgggctctccctctgccc 660  
DB 601 ctgctcgaggtgctgaggtctgctgcgagctgctgcgggctctccctctgccc 660  
QY 661 cgccgttccctcaactttcacacaataaagagagagccctctgtctatagttc 720  
DB 661 cgccgttccctcaactttcacacaataaagagagagccctctgtctatagttc 720

QY 721 ttatattcatcgaaaattacacataattacataagttatctgcttaaaaggagta 780  
DB 721 ttatattcatcgaaaattacacataattacataagttatctgcttaaaaggagta 780  
QY 781 attgagccgcaaaaggagtaattgagccgcaaaaggagtaattgagccgcaaaaggag 840  
DB 781 attgagccgcaaaaggagtaattgagccgcaaaaggagtaattgagccgcaaaaggag 840  
QY 841 gtaattgggcccagatcggtgtgtttacatggggaggaatcccttatcttcccca 900  
DB 841 gtaattgggcccagatcggtgtgtttacatggggaggaatcccttatcttcccca 900  
QY 901 ttgggaagacaacaagaatggccgacagccggcttcgacagacaaaatgtgtctc 960  
DB 901 ttgggaagacaacaagaatggccgacagccggcttcgacagacaaaatgtgtctc 960  
QY 961 ctgcgcaggtgagcgagaggggtctatatgtgcgacatcccccgcgcgcgcgtcaagc 1020  
DB 961 ctgcgcaggtgagcgagaggggtctatatgtgcgacatcccccgcgcgcgcgtcaagc 1020  
QY 1021 tcatgcatcctaataatgacacatcgcgcgccgcatgctgaatgagtgtgcgcatgaaa 1080  
DB 1021 tcatgcatcctaataatgacacatcgcgcgccgcatgctgaatgagtgtgcgcatgaaa 1080  
QY 1081 ttgcgcgtgagccgaatcgcgcaatcgacggcgatgaaaaccatgacgtgagagcctga 1140  
DB 1081 ttgcgcgtgagccgaatcgcgcaatcgacggcgatgaaaaccatgacgtgagagcctga 1140  
QY 1141 ccccgctgttcgaaggaactgaacgctcgcgtgtgtgacccaatgataccctgcgaagaatga 1200  
DB 1141 ccccgctgttcgaaggaactgaacgctcgcgtgtgtgacccaatgataccctgcgaagaatga 1200  
QY 1201 tctgtacagctcgcggtctgtgtcgaatgagcggaatagactaccgcgaagggcgagcg 1260  
DB 1201 tctgtacagctcgcggtctgtgtcgaatgagcggaatagactaccgcgaagggcgagcg 1260  
QY 1261 ggcgaactcctgtgaacgtggaaccttcgcgaggtatcatccgctgatacgcgagtgca 1320  
DB 1261 ggcgaactcctgtgaacgtggaaccttcgcgaggtatcatccgctgatacgcgagtgca 1320  
QY 1321 accactgggcccatttcgaacggtcaaacggtatccatctcgtgtagtattccgtgc 1380  
DB 1321 accactgggcccatttcgaacggtcaaacggtatccatctcgtgtagtattccgtgc 1380  
QY 1381 tgcgttccagacagctctcctgagctcgcgaatccttgatcagatgagcgcaaaaccttta 1440  
DB 1381 tgcgttccagacagctctcctgagctcgcgaatccttgatcagatgagcgcaaaaccttta 1440  
QY 1441 cgttcccgaggttgcggcgctcctcttgagtgcccgagggaaagatggttcgttgaaag 1500  
DB 1441 cgttcccgaggttgcggcgctcctcttgagtgcccgagggaaagatggttcgttgaaag 1500  
QY 1501 acgttaacagatttgcctcaaacctgcactgtagatgagatcaaacattatccgtctga 1560  
DB 1501 acgttaacagatttgcctcaaacctgcactgtagatgagatcaaacattatccgtctga 1560  
QY 1561 catlgtcgcaagagcgacgaacagatgtgcgtgaagctgtgcaagttgactaaggctgg 1620  
DB 1561 catlgtcgcaagagcgacgaacagatgtgcgtgaagctgtgcaagttgactaaggctgg 1620  
QY 1621 aagtgaagaagcgaacccaacgctgcgcagcgagctgaggttcccaaggtcgtgtcga 1680  
DB 1621 aagtgaagaagcgaacccaacgctgcgcagcgagctgaggttcccaaggtcgtgtcga 1680  
QY 1681 atgctgtcgcaagagggcgagcggaagatgagccccctctcccaagagcgcgga 1740  
DB 1681 atgctgtcgcaagagggcgagcggaagatgagccccctctcccaagagcgcgga 1740  
QY 1741 tcaactcaatccacgcttgcgtggaactgaaagcgtctgctgcagcaacaagagacaag 1800  
DB 1741 tcaactcaatccacgcttgcgtggaactgaaagcgtctgctgcagcaacaagagacaag 1800







QY	5101	ctctgagcgagagaacaaatacgcgaltcgcgtttaaaggacaattacaacaggaatcggat	5160
Db	5101	ccctgagcgagagaacgaatatacgatcgcgtttaaaggacaattacaacaggaatcggat	5160
QY	5161	gcaacccgcgcaggaacactctccacgcgatataaacattattccaaccigatacgaatat	5220
Db	5161	gcaacccgcgcaggaacgaacactgcgcgcgataccaataattttcaaccigatacgaatat	5220
QY	5221	cttctaataccctggaatcgtcttcttcccgggatcgcagtggttgagttaaccatgcat	5280
Db	5221	cttctaataccctggaatgctgttcttccgggattcgcagtggttgagtaaacatgcat	5280
QY	5281	cagaggtacgcgataaataatgctttgagtctcggaagcgataaatctcgttaagcagttta	5340
Db	5281	cagaggtacgcgataaataatgctttgagtctcggaagcgataaatctcgttaagcagttta	5340
QY	5341	gtctgaaccatctcatctgttaacatcatcttgcaacgcgtacaccttgcgcatgtttgaagaa	5400
Db	5341	gtctgaaccatctcatctgttaacatcatcttgcaacgcgtacaccttgcgcatgtttgaagaa	5400
QY	5401	actctggtcgcatcgggtcttcccatatacatgataatgtctgcacactgattgccgacat	5460
Db	5401	actctggtcgcatcgggtcttcccatatacatgataatgtctgcacactgattgccgacat	5460
QY	5461	tatcgcgagaccattataccaatataaatacagatccatctgttggaatttatcgcgcc	5520
Db	5461	tatcgcgagaccattataccaatataaatacagatccatctgttggaatttatcgcgcc	5520
QY	5521	tcgagcaagaagcttcccggtttgaatgatggctacataaacocctgtatctgttaatgt	5580
Db	5521	tcgagcaagaagcttcccggtttgaatgatggctacataaacocctgtatctgttaatgt	5580
QY	5581	aagcagaacagtttatgtgtcatgtatgtatatatttatacttctgtgaagtgaacatcga	5640
Db	5581	aagcagaacagtttatgtgtcatgtatgtatatatttatacttctgtgaagtgaacatcga	5640
QY	5641	gatttgaagacacaagctgtgcttcccccoccccatatttgaagcatttatcaggtt	5700
Db	5641	gatttgaagacacaagctgtgcttcccccoccccatatttgaagcatttatcaggtt	5700
QY	5701	atgtcttcacagacgagatalacatatattgtaattgataaataaacaataaggggttc	5760
Db	5701	atgtcttcacagacgagatalacatatattgtaattgataaataaacaataaggggttc	5760
QY	5761	cgcgcacattcccccgaagagtcgcacgtcttaagaacacattatcatgaacat	5820
Db	5761	cgcgcacattcccccgaagagagtcgcacgtcttaagaacacattatcatgaacat	5820
QY	5821	taacctataaataatagcgatatacgaagagcccttcgctc	58859
Db	5821	taacctataaataatagcgatatacgaagagcccttcgctc	58859
RESULT 2			
AAH79219/c			
ID	AAH79219	standard; DNA; 3604 BP.	
XX	AAH79219;		
AC	20-NOV-2001 (first entry)		
DT	Human leukocyte alpha-2b interferon production related DNA #12.		
DE	Human leukocyte alpha-2b interferon production; medicine; ds.		
XX	Human; leukocyte alpha-2b interferon production; medicine; ds.		
OS	Unidentified.		
XX	R02165455-C1.		
PN	20-APR-2001.		
PD	23-NOV-1999; 99RU-0124276.		
XX			
PF			







Accession	Source	Position	Sequence	Length
QY 5126	gcyttaaagaagcaattacaacacggaatcgaaatgcaacggcgaggaacacgtgcag	5185		
Db 1806	gctgttAAAGGACAAATTTCACAAACAGGAATCGAAATGCAACGGGCGAGAACATGCGCAG	1747		
QY 5186	cgcatacaaatatttcaacctggaatcaagatattcttctaatacctggaatgcgtttt	5245		
Db 1746	CGCATCAACAAATATTTTTCACCTGCAATACAGAGATATTCTTCTAATACCTGGAATGCTGTTTT	1687		
QY 5246	ccccgggagtcgcagtgagtgaaaccaatcatcatcagagatgaagataaaatggttgat	5305		
Db 1686	CCCCGGGATGCGACATGGTGTAGTAAACCATCATCATCAGAGTACGAGATGAATATGCTTGAT	1627		
QY 5306	ggtcggaagagggatcaattccgttcagccagtttagctggaacatctcatctgtaacatc	5365		
Db 1626	GGTGCGAAGAGGCAATTAATTCCTGACGCCAGATTGTAGTCTGACCAATCTCATCTGTAAACATC	1567		
QY 5366	attggcaacgctacacttctgcacatglttcagaaacaactctggcgcatcgggcttccata	5425		
Db 1566	ATTGGCAACGGGTACTCTTGGCATGTTTCAAGAAACAACCTCTGGCGCATCGGGCTTCCATTA	1507		
QY 5426	caatcgataagattgtgcgaactgattggccgacattatcggagcccaattataccata	5485		
Db 1506	CAATCGAATGAGATTGTGCACTGTATTGCCGACATTATGCGAGGCCAATTATACCCATA	1447		
QY 5486	taaatcagcatccatgltgtgaattaaatcggcgctcggagcaagacgtttcccgltgaat	5545		
Db 1446	TAAATCAGATPCATGTTGGAATTTAATCCGCGCTCGAGCAAGAGGTTTCCCGTTGAT	1387		
QY 5546	atggctcaataaaccccttgratracgtttagtaagaagacagatlttatgttcatga	5605		
Db 1386	ATGGCTCAATAAANCCCCCTGTATTACTGTTTATGTATGAACGACAGACAGTTTATTGTTCATGA	1327		
QY 5606	tgaatatatttatcttctgtgcaatgtaacatcagaagatttgagacaacagtgccttc	5665		
Db 1326	TGATATATTTTTATCTGTGCAATGTAAATCAAGAGATTTTGGAGACAACAGTGGCTTTC	1267		
QY 5666	ccccccccccattatggaagattatcagagttatgctcatagtcagcgatacatatt	5725		
Db 1266	CCCCCCCCCCCCATTATGGAAGCATTTATCAGGGTTATTGTGCTCATGACGCGAATACATATT	1207		
QY 5726	tgaatgatttagaanaataaacaacaa 5751			
Db 1206	TGAATGATTTTGAAGAAATAAACAACA 1181			
RESULT 3				
AAD09428				
ID	AAD09428 standard; DNA; 4867 BP.			
XX				
AC	AAD09428:			
XX				
DT	10-SEP-2001 (first entry)			
XX				
DE	Vaccine vector VJns.			
XX				
KM	Human immunodeficiency virus; HIV-1 Pol protein; VJ vector; vaccine;			
XX	codon optimisation; anti-HIV; immune response; HIV infection; bovine;			
XX	CMV; cytomegalovirus; ds.			
XX				
OS	Chimeric - Bos sp.			
OS	Chimeric - Human cytomegalovirus.			
OS	Chimeric - Unidentified.			
XX				
PN	WO200145748-A1.			
XX				
PD	28-JUN-2001.			
XX				
PF	21-DEC-2000; 2000WO-US34724.			
XX				
PR	22-DEC-1999; 99US-0171542.			
XX				
PA	(MER1) MERCK & CO INC.			
XX				

[illegible]



|||||  
Db 3041 taagctacactagaagaacagatatttgtaactgcctctgtgaagccattactctc 3100  
Oy 4226 gaaaaaaggttgtagccttgatccgcgcacaacacccgcctgtgagcggtgttt 4285  
Db 3101 gaaaaaaggttgtagccttgatccgcgcacaacacccgcctgtgagcggtgttt 3160  
Oy 4286 ttgtttgcaagcagcagatctacgcgcagaaaaaaggaatctcaagaagatcccttgatc 4345  
Db 3161 ttgtttgcaagcagcagatctacgcgcagaaaaaaggaatctcaagaagatcccttgatc 3220  
Oy 4346 ttcttcacggggtctgacgcctcagtggaacgaacactcacgttaaggatcttgatc 4405  
Db 3221 ttcttcacggggtctgacgcctcagtggaacgaacactcacgttaaggatcttgatc 3280  
Oy 4406 agatttcacaaaaaggaatcttcacactagatccctt ----- 4440  
Db 3281 agatttcacaaaaaggaatcttcacactagatccctttaaataatgaagttaaatca 3340  
Oy 4441 ----- 4440  
Db 3341 atctaagtatataatgataaacttgctgcaggttaccaatgcttaacagtgagca 3400  
Oy 4441 -----gggggggggggg 4452  
Db 3401 cctatctcagcgaatcgtctattctgtlcaatcgaagtgcctgactcgggggggggg 3460  
Oy 4453 cgttgtaggtctgctcgttgtagaagagtggtgctgactacacagagcttaacgcacca 4512  
Db 3461 cgttgtaggtctgctcgttgtagaagagtggtgctgactacacagagcttaacgcacca 3520  
Oy 4513 tcatcaagccagaagaagtgaaggagccacggctgtagagagcttggtgtagtgagcaag 4572  
Db 3521 tcatcaagccagaagaagtgaaggagccacggctgtagagagcttggtgtagtgagcaag 3580  
Oy 4573 ttggtgattttgaaactttgcttgccacggaaacggtcgtgctgcgggaagaatgctgt 4632  
Db 3581 ttggtgattttgaaactttgcttgccacggaaacggtcgtgctgcgggaagaatgctgt 3640  
Oy 4633 atctgaccccttcaactcagcaaaaggtcgaatttaacaacaaacgcgcgtccgcgaag 4692  
Db 3641 atctgaccccttcaactcagcaaaaggtcgaatttaacaacaaacgcgcgtccgcgaag 3700  
Oy 4693 tcaagcgaatgctcgcagtggtatacaaccaatlaaccaatctgcatagaaaaactcat 4752  
Db 3701 tcaagcgaatgctcgcagtggtatacaaccaatlaaccaatctgcatagaaaaactcat 3760  
Oy 4753 cgagcatcaaatgaactcgaatttatcataatcagagattacaatacatattttgaa 4812  
Db 3761 cgagcatcaaatgaactcgaatttatcataatcagagattacaatacatattttgaa 3820  
Oy 4813 aaagccgttctgtaatgaagaagaaaaactcagcgaagcaattcataggaatggcaagat 4872  
Db 3821 aaagccgttctgtaatgaagaagaaaaactcagcgaagcaattcataggaatggcaagat 3880  
Oy 4873 cctgtatcgggtcgtcgaattccgaactcgtccaacatcaatacaacttaattccctc 4932  
Db 3881 cctgtatcgggtcgtcgaattccgaactcgtccaacatcaatacaacttaattccctc 3940  
Oy 4933 cgtcaaaaataaggttatcaagtgagaataatcacatgagtgaagctaaatccggtgga 4992  
Db 3941 cgtcaaaaataaggttatcaagtgagaataatcacatgagtgaagctaaatccggtgga 4000  
Oy 4993 atggcaaaaaggttatgcaattctcttcagagctgttcaacaggccagcatlaagctgt 5052  
Db 4001 atggcaaaaaggttatgcaattctcttcagagctgttcaacaggccagcatlaagctgt 4060  
Oy 5053 catcaaaaatcaactcgcacatcaaaccaacggttatcatcatcgtgattgcgctgagcgagac 5112  
Db 4061 catcaaaaatcaactcgcacatcaaaccaacggttatcatcatcgtgattgcgctgagcgagac 4120  
Oy 5113 gaatacgcggtcgtgttaaaaggaataatcaaacaggaatcgaaatgcaacccggcgca 5172  
|||||

Db 4121 gaatacgcggtcgtgttaaaaggaataatlaacaacaggaaatcgaaatgcaacccggcgca 4180  
Oy 5173 ggaacacgtgcagcgaatcaacaatatctacactgaataagaatattcttcaactc 5232  
Db 4181 ggaacacgtgcagcgaatcaacaatatctacactgaataagaatattcttcaactc 4240  
Oy 5233 ggaatcgtttctccggggaatcgagtggttgtaacatgataatcaaggagtaagga 5292  
Db 4241 ggaatcgtttctccggggaatcgagtggttgtaacatgataatcaaggagtaagga 4300  
Oy 5293 taaatgcttgatggttcgggaaggagcaataatccgtcagccagttatgactaaccatc 5352  
Db 4301 taaatgcttgatggttcgggaaggagcaataatccgtcagccagttatgactaaccatc 4360  
Oy 5353 catctgtaacatcatltygcaacgctacacttgccatgttccaagaacactcggcgcat 5412  
Db 4361 catctgtaacatcatltygcaacgctacacttgccatgttccaagaacactcggcgcat 4420  
Oy 5413 cgggcttcccatcaaatcgaatagatgtgcacactgattgcccgaacattatcgagacc 5472  
Db 4421 cgggcttcccatcaaatcgaatagatgtgcacactgattgcccgaacattatcgagacc 4480  
Oy 5473 attatacccatataaatacgaatcccatgttggaaattaatcgcgcctcgaagaagcg 5532  
Db 4481 attatacccatataaatacgaatcccatgttggaaattaatcgcgcctcgaagaagcg 4540  
Oy 5533 ttcccgcttgaaatagtgctcaataaacaccccttgattactgatttaatgaaagacagct 5592  
Db 4541 ttcccgcttgaaatagtgctcaataaacaccccttgattactgatttaatgaaagacagct 4600  
Oy 5593 ttattgtlcatagatataatttatctgttgtaagtgaacatcaagaatatttgagaca 5652  
Db 4601 ttattgtlcatagatataatttatctgttgtaagtgaacatcaagaatatttgagaca 4660  
Oy 5653 caacgtgcttcccccctcccatattatgaacatattatcagaggtatgtctcagba 5712  
Db 4661 caacgtgcttcccccctcccatattatgaacatattatcagaggtatgtctcagba 4720  
Oy 5713 gcggaatacatattgaaatgatttaagaanaataaacaatagaggtccgcgacattc 5772  
Db 4721 gcggaatacatattgaaatgatttaagaanaataaacaatagaggtccgcgacattc 4780  
Oy 5773 cccgaaaaagtgccacgtcgtcttaagaacatattatcagaggtatgtctcagba 5832  
Db 4781 cccgaaaaagtgccacgtcgtcttaagaacatattatcagaggtatgtctcagba 4840  
Oy 5833 atagcgatcacgagcccttcgct 5859  
Db 4841 atagcgatcacgagcccttcgct 4867  
  
RESULT 4  
AAD09614  
ID AAD09614 standard; DNA; 4867 BP.  
XX  
XX AAD09614;  
DT 10-SEP-2001 (first entry)  
XX  
XX  
DE Vaccine vector VJUNs.  
XX  
XX DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
KW cell mediated immune response; cytotoxic T lymphocyte; CTL;  
KW human immunodeficiency virus; HIV; vaccine vector; ds.  
XX  
OS Chimeric - Human cytomagalovirus.  
OS Chimeric - Bos sp.  
OS Chimeric - Unidentified.  
PN  
XX  
XX WO200143693-A2.  
XX  
XX 21-JUN-2001.  
XX



QY	4046	cgcttgatgtccaaacccggtlaagacacagatlaacgcacactgacgacgacactgttaac	4105
Db	2921	cgctctgattccaaaccccggtlaagacacgacttaacgcactgacgacgacactgttaac	2960
QY	4106	agattgacgagacgagatgtatgtaggcggtgtctacagagttcttgaagtgtgtgctaac	4165
Db	2981	agattgacgagacgagatgtatgtaggcggtgtctacagagttcttgaagtgtgtgctaac	3040
QY	4166	tacggtcacactagaaggaacagatattgtatctgcgtctgcgtgaagccagttacttc	4225
Db	3041	tacggtcacactagaaggaacagatattgtatctgcgtctgcgtgaagccagttacttc	3100
QY	4226	ggaagaaagtgtgtagcctcttgatccgcgaacaacaccgcgctgtgtagcgtgttt	4285
Db	3101	ggaagaaagtgtgtagcctcttgatccgcggaacaacaccgcgctgtgtagcgtgttt	3160
QY	4286	ttgtgttcgaagacagacagatlaacgcgcgagaaaaaagatctcaagaagatcccttgatc	4345
Db	3161	ttgtgttcgaagacagacagatlaacgcgcgagaaaaaagatctcaagaagatcccttgatc	3220
QY	4346	ttttctacggggtcttgagcgtcagtgtggaacgaaaaactacglttaagagatttggatcatg	4405
Db	3221	ttttctacggggtcttgagcgtcagtgtggaacgaaaaactacglttaagagatttggatcatg	3280
QY	4406	agattatcaaaaaagatcttcaaccatagatctttt-----	4440
Db	3281	agattatcaaaaaagatcttcaaccatagatcttttlaaataatgaagttttaatca	3340
QY	4441	-----	4440
Db	3341	atctaaagatataatgatlaaactgtgtcgtacagttaccatgtctaaacgltgaagca	3400
QY	4441	-----	4452
Db	3401	ccctactcagcagatctgtctcttccgttcaatccatagttgcttgactcggggggggggg	3460
QY	4453	cgctgaagctctgcgctcgttgaagaagagtgctgtgctgactacataccagacgtgaatgccca	4512
Db	3461	cgctgaagctctgcgctcgttgaagaagagtgctgtgctgactacataccagacgtgaatgccca	3520
QY	4513	tcaatccacccgaagaagtgaagggaaacccaggttataagaagccttgtgttaggtgacacag	4572
Db	3521	tcaatccacccgaagaagtgaagggaaacccaggttataagaagccttgtgttaggtgacacag	3580
QY	4573	ttgtgtgatttgaactttgctcttgccacggaacggtctgcgctgtgcgggaagatgcgtg	4632
Db	3581	ttgtgtgatttgaactttgctcttgccacggaacggtctgcgctgtgcgggaagatgcgtg	3640
QY	4633	atctgataccttcaacactacagaagaagtgtgatttatcttaacaaagccgcgcgtccgctcaag	4692
Db	3641	atctgataccttcaacactacagaagaagtgtgatttatcttaacaaagccgcgcgtccgctcaag	3700
QY	4693	tcagcgtaatgtcttcgacgtgttataacacaaattaaaccaattctgtattagaanaaacat	4752
Db	3701	tcagcgtaatgtcttcgacgtgttataacacaaattaaaccaattctgtattagaanaaacat	3760
QY	4753	cgagcatcaaaatgaactctgaatttatataatcagaagtatcaatacatatatttgaa	4812
Db	3761	cgagcatcaaaatgaactctgaatttatataatcagaagtatcaatacatatatttgaa	3820
QY	4813	aaagccgcttctctgtaatgaaggagaanaactacccgagacagttccatagaatgcaagat	4872
Db	3821	aaagccgcttctctgtaatgaaggagaanaactacccgagacagttccatagaatgcaagat	3880
QY	4873	ccctggtatccggtctgcgattccgactcgttcaacaatcaatacaaccttaattccoct	4932
Db	3881	ccctggtatccggtctgcgattccgactcgttcaacaatcaatacaaccttaattccoct	3940
QY	4933	cgctcaaaaataaggttatataagttggaatacaatagatgtagcagactggaatccggtgaga	4992
Db	3941	cgctcaaaaataaggttatataagttggaatacaatagatgtagcagactggaatccggtgaga	4000







Db 2819 cctgtccgacccgtcgcttaaccggaatacctgtccgccttctcccttcgggaagctg 2878  
Qy 3926 ggcgttttcctagctcagctcgtgttagatatactcaagttccggtgtgctgtccgaag 3985  
Db 2879 ggcgttttcctagctcagctcgtgttagatatacctcgtgtgctgtccgaag 2938  
Qy 3986 ctggagctgtgtgcaagaaaccccggttccagccgacgctgcgcttatccggaactat 4045  
Db 2939 ctggagctgtgtgcaagaaaccccggttccagccgacgctgcgcttatccggaactat 2998  
Qy 4046 cgtcttgagttccaaacccggttaagaaacgactatccgcaactggaagccactgtgtaac 4105  
Db 2999 cgtcttgagttccaaacccggttaagaaacgactatccgcaactggaagccactgtgtaac 3058  
Qy 4106 aggattagcagagcgaaggtatgttaggcgtgtcagagatctctgaaagtgtgcccatac 4165  
Db 3059 aggattagcagagcgaaggtatgttaggcgtgtcagagatctctgaaagtgtgcccatac 3118  
Qy 4166 taggctcactatagaagaacagatattgtgtatctgcgctcgtcgtcgaagccaattacctc 4225  
Db 3119 taggctcactatagaagaacagatattgtgtatctgcgctcgtcgtcgaagccaattacctc 3178  
Qy 4226 ggaanaaagatctgttagctctgtatccgcgaacaacacccgcgtgtgtagcgtgtttt 4285  
Db 3179 ggaanaaagatctgttagctctgtatccgcgaacaacacccgcgtgtgtagcgtgtttt 3238  
Qy 4286 ttgtttgcaagcgcagcagattacgcgcagaanaaagatcctcaagaagatccttgatc 4345  
Db 3239 ttgtttgcaagcgcagcagattacgcgcagaanaaagatcctcaagaagatccttgatc 3298  
Qy 4346 ttctctaagggtctgacgctcagctggaacgaaacacacgcttaagagatttggatc 4405  
Db 3299 ttctctaagggtctgacgctcagctggaacgaaacacacgcttaagagatttggatc 3358  
Qy 4406 agattatcaaaaagatcctcactagatccttt----- 4440  
Db 3359 agattatcaaaaagatcctcactagatccttttaataaataagatttaataca 3418  
Qy 4441 ----- 4440  
Db 3419 attcaaaagtatatatgtgtaacttggctcgtacaggttaaccaatgcttaacgttaggca 3478  
Qy 4441 -----ggggggggggggg 4452  
Db 3479 ccatactcagcagatcgtctatacttccgttcaaccatagttgcgtaactcgggggggggg 3538  
Qy 4453 cgcgtgaggtctgcctcgttgaagaaggtgtgtgcacatcacagggccttgaatccgccca 4512  
Db 3539 cgcgtgaggtctgcctcgttgaagaaggtgtgtgcacatcacagggccttgaatccgccca 3598  
Qy 4513 tcaatccagcgaagaagtgaaggagccagcgtgtgtagagccttctgttaggtgagccag 4572  
Db 3599 tcaatccagcgaagaagtgaaggagccagcgtgtgtagagccttctgttaggtgagccag 3658  
Qy 4573 ttggtgatttgaactttctgtccagcgaacggtcgtcgtgtgcgggaagaatgcgtg 4632  
Db 3659 ttggtgatttgaactttctgtccagcgaacggtcgtcgtgtgcgggaagaatgcgtg 3718  
Qy 4633 attcgtatccttcaactcagcaaaaagtgcgaattatccaacaagccgcgcgtcccgtaag 4692  
Db 3719 attcgtatccttcaactcagcaaaaagtgcgaattatccaacaagccgcgcgtcccgtaag 3778  
Qy 4693 tcaacgtatcgtctgcaggtgttacaacaacatcaacatcttctgtttagaanaactat 4752  
Db 3779 tcaacgtatcgtctgcaggtgttacaacaacatcaacatcttctgtttagaanaactat 3838  
Qy 4753 cgaagcatcaaaatgaactgcaattatcatalcagagatlatcaatacatalttttgaa 4812  
Db 3839 cgaagcatcaaaatgaactgcaattatcatalcagagatlatcaatacatalttttgaa 3898  
Qy 4813 aaagccgttctcttaagaaggagaaacacccagcagcaggttccatagatgagcaagat 4872  
Db 3899 aaagccgttctcttaagaaggagaaacacccagcagcaggttccatagatgagcaagat 3958

Qy 4873 cctgtatcgtctcgtgaltccgactcgtlccaacalcaatacaacttaattccct 4932  
Db 3959 cctgtatcgtctcgtgaltccgactcgtlccaacalcaatacaacttaattccct 4018  
Qy 4933 cgtcaaaaataaggtatcaaaatgagaaatcccatctgtgtagagatccgttga 4992  
Db 4019 cgtcaaaaataaggtatcaaaatgagaaatcccatctgtgtagagatccgttga 4078  
Qy 4993 atggcaaaagcttatgcatcttcttccagactgtgtcaacagcgcacatccatccgt 5052  
Db 4079 atggcaaaagcttatgcatcttcttccagactgtgtcaacagcgcacatccatccgt 4138  
Qy 5053 catcaaaatcactcgcataccaacacccgttatcatctcgtgattgcgcttgaagcagac 5112  
Db 4139 catcaaaatcactcgcataccaacacccgttatcatctcgtgattgcgcttgaagcagac 4198  
Qy 5113 gaaatacgcgatacgtgtttaaagaagacaattacaacacagatcgaatgcaacgcgca 5172  
Db 4199 gaaatacgcgatacgtgtttaaagaagacaattacaacacagatcgaatgcaacgcgca 4258  
Qy 5173 ggaacactcgcagcgcataccaatatttcaactgtaacgttagatcttataact 5232  
Db 4259 ggaacactcgcagcgcataccaatatttcaactgtaacgttagatcttataact 4318  
Qy 5233 ggaatgctgtttcccgggagtcgagtggtgagtaaacatgcatcaacgagtagcga 5292  
Db 4319 ggaatgctgtttcccgggagtcgagtggtgagtaaacatgcatcaacgagtagcga 4378  
Qy 5293 taaatgctgtgtagtgtaggaaggcagataaattccgtlcaagcgaattagctgaacct 5352  
Db 4379 taaatgctgtgtagtgtaggaaggcagataaattccgtlcaagcgaattagctgaacct 4438  
Qy 5353 catctgtatacatctgtgcaacgcgtacacttgcagattttaaagaanaactctgcgcat 5412  
Db 4439 catctgtatacatctgtgcaacgcgtacacttgcagattttaaagaanaactctgcgcat 4498  
Qy 5413 cgggtctcccatcaatcgtatagattgtgcacactgtatgcgcgaattatcgcgagcc 5472  
Db 4499 cgggtctcccatcaatcgtatagattgtgcacactgtatgcgcgaattatcgcgagcc 4558  
Qy 5473 attatacccatataaatacgaatccatcgtgtggaatttaacgcggcctcgaagaagc 5532  
Db 4559 attatacccatataaatacgaatccatcgtgtggaatttaacgcggcctcgaagaagc 4618  
Qy 5533 ttcccggttgaatatgtgtcataacaccccttgaattactgtttagtgaagcagaagt 5592  
Db 4619 ttcccggttgaatatgtgtcataacaccccttgaattactgtttagtgaagcagaagt 4678  
Qy 5593 ttatgttcaatgataataatttatactgtgtgcaatgataacacagaattttgagaca 5652  
Db 4679 ttatgttcaatgataataatttatactgtgtgcaatgataacacagaattttgagaca 4738  
Qy 5653 caagctggtcttcccccccccccatattgaagacattatcaagggttatgttctcatga 5712  
Db 4739 caagctggtcttcccccccccccatattgaagacattatcaagggttatgttctcatga 4798  
Qy 5713 gcgatacataattggaatgttattagaanaaataaacaataagggttccgcgcacattc 5772  
Db 4799 gcgatacataattggaatgttattagaanaaataaacaataagggttccgcgcacattc 4858  
Qy 5773 cccgaaaaatgcccacccgtgaagcgttcaagaacacatlatatcatgacatlaacctataaa 5832  
Db 4859 cccgaaaaatgcccacccgtgaagcgttcaagaacacatlatatcatgacatlaacctataaa 4918  
Qy 5833 ataggctatcagagcccttccgtc 5859  
Db 4919 ataggctatcagagcccttccgtc 4945

RESULT 6  
AAD09607  
ID AAD09607 standard; DNA; 4945 BP.



AC AAD09607;  
 XX 10-SEP-2001 (first entry)  
 XX  
 DE VJLns-tPA vector to construct human tPA-HIV-1 nef DNA fusion construct.  
 XX  
 XX  
 XX DNA vaccine; codon optimisation; Nef protein; HIV infection; anti-HIV;  
 XX cell mediated immune response; cytotoxic T lymphocyte; CTL;  
 KW human immunodeficiency virus-1; HIV-1; VJLns-tPA vector; ds.  
 XX  
 OS Chimeric - Human immunodeficiency virus type 1.  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FH misc\_feature 1877..1960  
 FT /\*tag= a  
 FT /note= "Human tPA leader sequence encoding DNA"  
 FT  
 XX WO200143693-A2.  
 PN  
 XX  
 PD 21-JUN-2001.  
 PD  
 PE 15-DEC-2000; 2000WO-US4162.  
 XX  
 XX 17-DEC-1999; 99US-0172442.  
 PR  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX Shiver JW, Liang X, Fu T;  
 PL  
 DR WPI: 2001-417878/44.  
 XX  
 XX Novel DNA vaccine useful for inducing cell mediated immune response  
 PT against virulent strains of HIV, comprises DNA expression vector, and  
 PT DNA containing codon optimized open reading frame encoding Nef protein  
 PT  
 PT  
 XX  
 XX Example 1; Page 36-38; 84pp; English.  
 XX  
 XX  
 XX The present invention relates to a DNA vaccine comprising a DNA  
 CC expression vector, and a DNA molecule containing a codon optimised open  
 CC reading frame encoding a Nef protein or its immunogenic derivative. The  
 CC Nef protein or its derivative is expressed and generates an immune  
 CC response which provides a substantial level of protection against HIV  
 CC infection, upon administration of to a host. The DNA vaccine is useful  
 CC for inducing a cell mediated immune (cytotoxic T lymphocyte, CTL)  
 CC response against infection or disease caused by virulent strains of HIV,  
 CC by administering the DNA vaccine into the tissue of the vertebrate host,  
 CC preferably human. It is also useful for lowering transmission rate to  
 CC previously uninfected individuals and/or for reducing levels of viral  
 CC loads within an infected individual, so as to prolong the asymptomatic  
 CC phase of HIV infection. The present sequence is a vaccine vector  
 CC VJLns-tPA to construct human tissue-specific plasminogen activator (tPA)  
 CC and HIV-1 pol DNA fusion construct.  
 CC  
 XX Sequence 4945 BP; 1245 A; 1244 C; 1199 G; 1257 T; 0 other;  
 XX

	Query Match	Best Local Similarity	Score	DB	Length
Matches 2293;	Conservative	0;	Mismatches	1;	Indels 133; Gaps 1;
0Y	3566	gcttcctcgctacagcagctcgctgctgcgttcgctcgctcgcggaacggtatcagct	3625		
Db	2519	gcttcctcgctacagcagctcgctgcgttcgctgcgtgcggaacggtatcagct	2578		
0Y	3626	cactcaaaagcggtatatacgtgtatccacagatcagggataaacgcagaagaacatg	3685		
Db	2579	cactcaaaagcggtatatacgtgtatccacagatcagggataaacgcagaagaacatg	2638		
0Y	3686	tgaagcaaaagcgacgaacaaagcgacgaacglaaaaaagccgcggttcgttcgcttttc	3745		

Db	2639	cgagcaaaagccagcaaaaagccaaagaaacgcttaaaaaggccgcgtgtgcagcgcttttc	2698
Oy	3746	cgatagctccgcccccttgaacgagcaatcaaaaaatcgaagctcaaaatcgaagtgtgcga	3805
Db	2699	catagagctccgcccccttgaacgagcaatcaaaaaatcgaagctcaaaatcgaagtgtgcga	2758
Oy	3806	aaaccgacagactataaagataccaagcgctttcccccttgaaagctccctctgagctct	3865
Db	2759	aaaccgacagactataaagataccaagcgctttcccccttgaaagctccctctgagctct	2818
Oy	3866	ccgtgtccgaacctgcccgtctaacggaataactgtccgccttttcccttcgggaagcggtg	3925
Db	2819	ccgtgtccgaacctgcccgtctaacggaataactgtccgccttttcccttcgggaagcggtg	2878
Oy	3926	ggcccttcataagatgcacagcggttaagtaactaaatctcgggtgtagaagcggtctccaa	3985
Db	2879	ggcccttcataagatgcacagcggttaagtaactaaatctcgggtgtagaagcggtctccaa	2938
Oy	3986	ctggagctgtgtgacagcaaacccccgtctacagcccgagccgctgtgcctatcccgtaaat	4045
Db	2939	ctggagctgtgtgacagcaaacccccgtctacagcccgagccgctgtgcctatcccgtaaat	2998
Oy	4046	cgctcttgagttccaaaccccggttaagaacacgactatcgcacatgtgcagcagccactgtgtac	4105
Db	2999	cgctcttgagttccaaaccccggttaagaacacgactatcgcacatgtgcagcagccactgtgtac	3058
Oy	4106	aggatttagcagaagcagaggtatgtagaagcggtgcacagagttcttgaaggtgtgtgccaac	4165
Db	3059	aggatttagcagaagcagaggtatgtagaagcggtgcacagagttcttgaaggtgtgtgccaac	3118
Oy	4166	taacggtacacatgaagaagcagcaatcttggatctgtgcctctgtctgtgaagccagttaccttc	4225
Db	3119	taacggtacacatgaagaagcagcaatcttggatctgtgcctctgtctgtgtgaagccagttaccttc	3178
Oy	4226	ggaaaaaaggtgtgtgctcttgaatcccggaacaaacacacgcgtgtgtagcgtgtgtttt	4285
Db	3179	ggaaaaaaggtgtgtgctcttgaatcccggaacaaacacacgcgtgtgtgtagcgtgtgtttt	3238
Oy	4286	tttgtttgaagaagcagagatattaagcgcagagaaaaaagatctctaaagagatccctttatc	4345
Db	3239	tttgtttgaagaagcagagatattaagcgcagagaaaaaagatctctaaagagatccctttatc	3298
Oy	4346	ttttctacggggtctgaagctcagcttggaagaacacacagtttaagagattttgtcatgt	4405
Db	3299	ttttctacggggtctgaagctcagcttggaagaacacacagtttaagagattttgtcatgt	3358
Oy	4406	agattatcaaaaaagatctcacctagatccttt-----	4440
Db	3359	agattatcaaaaaagatctcacctagatccttttaataatlaaaatgaagttttaatca	3418
Oy	4441	-----	4440
Db	3419	atctaaagatataatagtaaacctgtgtcgtacagttaccaaatgtcctaacagtgtgaagca	3478
Oy	4441	-----	4452
Db	3479	ccctattccagcgatctgtctatcttggttcatcctaagttgctccgacccgggggggggggg	3538
Oy	4453	cgcttgagctcgtcgtctgtaagaagaagttgtctgtgactacataccagcgttgaaatgcacca	4512
Db	3539	cgcttgagctcgtcgtctgtaagaagaagttgtctgtgactacataccagcgttgaaatgcacca	3598
Oy	4513	tcatccagccagaanaatgtgaaggggccaaggtgtatgtagaagactttgtgtatgtgtgaagca	4572
Db	3599	tcatccagccagaanaatgtgaaggggccaaggtgtatgtagaagactttgtgtatgtgtgaagca	3658
Oy	4573	ttggtatatttgaacttttgcttgccaaggaagcggtcgtgtgtctgcgggaagaatgtggtg	4632
Db	3659	ttggtatatttgaacttttgcttgccaaggaagcggtcgtgtgtctgcgggaagaatgtggtg	3718
Oy	4633	atctgtatccttcaactcagcaaaaagttcgaatttatcctaaacaaagccgcgtcccgctcaag	4692



OY	5773	ccgagaagtcgccaacgcgaagtcctaagaacacattatctatctgacataaaccttaaa	5832
Db	4859	cccgaaagtgcccccgcgaagtcctaaagaacacattatctatgacataaaccttaaa	4918
OY	5833	ataagcgtatccagagagccttcctgc	5859
Db	4919	ataagcgtatccagagagccttcctgc	4945

## AAC86258

XX  
AC AAC86258;

AC AAC86258;

DT 28-FEB-2001 (first entry)

DE Plasmid GHRH1-29WTCMV

KW GHRH; growth hormone-releasing hormone; ds.

OS Synthetic.

PN EP1052286-A2

PD 15-NOV-2000.

PF 12-APR-2000; 2000EP-0302790.

PR 12-APR-1999; 99US-0128830.

PA (PFIZ ) PFIZER PROD INC.

PI Morsey MA, Sheppard MG;

DR WPI; 2001-026585/04.

PT New polypeptide variants of growth hormone releasing hormone with

PT growth hormone deficiency related disorders or to improve growth and

XX

PS Examples; Page 39-41; 67pp; English.

CC The present invention relates to growth hormone-releasing hormone  
CC (GHRH) variants having enhanced resistance to enzymatic degradation.

CC treat growth hormone deficiency related disorders, or to improve

pharmaceutical compositions to promote expression and elevation of

higher levels than prior art variants modified using traditional

degradation, therefore have increased length of activity.

Sequence 5108 BP: 1279 A: 1312 C: 1231 G: 1286 T: 0 other:

Sequence 5108 BP; 1279 A; 1312 C; 1231 G; 1286 T; 0 other;

Query Match 36.78; Score 2149.4; DB 22; Length 5108;

Matches 2293; Conservative 0; Mismatches 1; Indels 133; Gaps 1;

QV 3566 qcttcctcactgactcagctgcgtcgaatcattcaggcttgcacggcgcaaacgataccaagt 3625

Db 554 acctcctcaactgaactcagtcagtccagctacgacgcgaaccgatatacact 613

3626 cactcaaaagcgaataacgttatcaccaaatcaaggaataaaacgaacgaagaaaagaaacatg 3688

Dh  
614 cactcaaaacatatacattatccaccagaatcacagataaacccaagaaaacaatg 673

Ov 3686 +aaacaaaancccaaaaaaaccaaaaacccaatataaaanccccccatttctttc 3745

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84



Db 674 tgaagcaaaagccagcaaaaagccaggaacgltaaaagccgcgtgtcgtgacgtttttc 733  
Qy 3746 cataagctccgccccctgacgagcatcacaaaaatcgacgtcaagtacagagtgagca 3805  
Db 734 cataggtccgcgccccctgacgagcatcacaaaaatcgacgtcaagtacagagtgagca 793  
Qy 3806 aaccgcagagactataaagatacagagcgtttccccctggaaagctccctcgtagctc 3865  
Db 794 aaccgcagagactataaagatacagagcgtttccccctggaaagctccctcgtagctc 853  
Qy 3866 cctgttcgcagccctgcgcgttacccgatacactgttcgccttcccttcggaaagcgtg 3925  
Db 854 cctgttcgcagccctgcgcgttacccgatacactgttcgccttcccttcggaaagcgtg 913  
Qy 3926 gcgccttcataagctacgcgtgtagtatacagltcggtgggttaagtcgttcgcgcag 3985  
Db 914 gcgccttcataagctacgcgtgtagtatacagltcggtgggttaagtcgttcgcgcag 973  
Qy 3986 ctgggctgtgtgcagaaaccccccttcagccccgcgcgtgcgcttalcgtaactat 4045  
Db 974 ctgggctgtgtgcagaaaccccccttcagccccgcgcgtgcgcttalcgtaactat 1033  
Qy 4046 cgtcttgatgccaaacccggttaagacaacgactatcgcacgtgcgcagccacgtgtaac 4105  
Db 1034 cgtcttgatgccaaacccggttaagacaacgactatcgcacgtgcgcagccacgtgtaac 1093  
Qy 4106 aggaattacagagcaggtatgtatgagcgtgtgtacagagcttcgaaagtggcctaac 4165  
Db 1094 aggaattacagagcaggtatgtatgagcgtgtgtacagagcttcgaaagtggcctaac 1153  
Qy 4166 taagctacactagaagaagcagatlttgatctgcgtcgtcgtgtaagccagttaccttc 4225  
Db 1154 taagctacactagaagaagcagatlttgatctgcgtcgtcgtgtaagccagttaccttc 1213  
Qy 4226 ggaaaaaagttgtgagctcttgatccgcgcaaaaacacacccgcgtgtagcggtgttt 4285  
Db 1214 ggaaaaaagttgtgagctcttgatccgcgcaaaaacacacccgcgtgtagcggtgttt 1273  
Qy 4286 ttgtgttcagagcagatcacgcgcgcaaaaaggaatcccaagaagatcccttgatc 4345  
Db 1274 ttgtgttcagagcagatcacgcgcgcaaaaaggaatcccaagaagatcccttgatc 1333  
Qy 4346 ttcttaacgggtcttcgacgtcagtggaacgaaacacgcgttaagggatttgltcatg 4405  
Db 1334 ttcttaacgggtcttcgacgtcagtggaacgaaacacgcgttaagggatttgltcatg 1393  
Qy 4406 agattatacaaaaagatcttcacatcctt----- 4440  
Db 1394 agattatacaaaaagatcttcacatcctttaaataaagaagtttaataca 1453  
Qy 4441 ----- 4440  
Db 1454 atctaaatataatgagtaaaactcgtcgtacagltaccaaagtcttaacgtttaagca 1513  
Qy 4441 -----ggggggggggggg 4452  
Db 1514 cctaatcagcgatcgtctatcttcgttcacatcagttgctgactcggggggggggg 1573  
Qy 4453 cgtgtgaagtcgtccctcgtgaagaaggtgtgtgtgactataccaaagcctcgaaatcgccca 4512  
Db 1574 cgtgtgaagtcgtccctcgtgaagaaggtgtgtgtgactataccaaagcctcgaaatcgccca 1633  
Qy 4513 tcaatccacgaagaagtgagggagccacggttgatgagagcttctgtgagtgagccag 4572  
Db 1634 tcaatccacgaagaagtgagggagccacggttgatgagagcttctgtgagtgagccag 1693  
Qy 4573 ttgtgtatttgaaactttgcttgcgtccaggaacggtcgtcgtcgtcgggaagatgctgtg 4632  
Db 1694 ttgtgtatttgaaactttgcttgcgtccaggaacggtcgtcgtcgtcgggaagatgctgtg 1753  
Qy 4633 atctgatacttcaactcagcaaaagttcgatttatccaacaagccgcgtccgtccaag 4692  
Db 1754 atctgatacttcaactcagcaaaagttcgatttatccaacaagccgcgtccgtccaag 1813

Qy 4693 tcaagctaatgctctgcagtglttaacaaccaatlaaccaatctgattagaanaactcat 4752  
Db 1814 tcaagctaatgctctgcagtglttaacaaccaatlaaccaatctgattagaanaactcat 1873  
Qy 4753 cgaagcatcaaatgaaactgaaattatcatatcaggaattacatatacatatttgaa 4812  
Db 1874 cgaagcatcaaatgaaactgaaattatcatatcaggaattacatatacatatttgaa 1933  
Qy 4813 aaagccgtttcgttaatgaaaggaanaactcagagaggttccaatgagtgagcaagt 4872  
Db 1934 aaagccgtttcgttaatgaaaggaanaactcagagaggttccaatgagtgagcaagt 1993  
Qy 4873 cctgtatcgtctgcagatccgactcgtccaaactcaataacatataatcttccct 4932  
Db 1994 cctgtatcgtctgcagatccgactcgtccaaactcaataacatataatcttccct 2053  
Qy 4933 cgtcaaaaaaaggtatcatcaagtgagaaatcacatgagtgagacgtgaatccggtgaga 4992  
Db 2054 cgtcaaaaaaaggtatcatcaagtgagaaatcacatgagtgagacgtgaatccggtgaga 2113  
Qy 4993 atgcaaaaaggtatgcatcttcttcagactgtgtccaacagccagccatgacgtgt 5052  
Db 2114 atgcaaaaaggtatgcatcttcttcagactgtgtccaacagccagccatgacgtgt 2173  
Qy 5053 catcaaaatcactcgcataccaacaaacgltatcatcgttatgtgcgtgagcgagac 5112  
Db 2174 catcaaaatcactcgcataccaacaaacgltatcatcgttatgtgcgtgagcgagac 2233  
Qy 5113 gaaatacgcgactcgtgttaaaaggaacatctcaaaaaggaatcgaaatgcaacccgagca 5172  
Db 2234 gaaatacgcgactcgtgttaaaaggaacatctcaaaaaggaatcgaaatgcaacccgagca 2293  
Qy 5173 ggaacactgcgcgcgcatcaaatatcttcacgtgaatcaagatatttctaactc 5232  
Db 2294 ggaacactgcgcgcgcatcaaatatcttcacgtgaatcaagatatttctaactc 2353  
Qy 5233 ggaatgctgtttcccgagatcgcagtggtgagtaacatcgcatacaggaagtagcaga 5292  
Db 2354 ggaatgctgtttcccgagatcgcagtggtgagtaacatcgcatacaggaagtagcaga 2413  
Qy 5293 taaaatgcttgaatgcggaagaggaataatccgtcagcaggtttagtctgacaact 5352  
Db 2414 taaaatgcttgaatgcggaagaggaataatccgtcagcaggtttagtctgacaact 2473  
Qy 5353 catctgttaacatcatgttggaacgcgtacactttgcaggtttcagaanaactctgggcact 5412  
Db 2474 catctgttaacatcatgttggaacgcgtacactttgcaggtttcagaanaactctgggcact 2533  
Qy 5413 cgggcttcccatacaatcgatagatgtgcgacactgattgcgcgacatatacgcgagcc 5472  
Db 2534 cgggcttcccatacaatcgatagatgtgcgacactgattgcgcgacatatacgcgagcc 2593  
Qy 5473 attatataccatataaatacagcatccaatgtcgaaatttaacogcgcctcagaaagcg 5532  
Db 2594 attatataccatataaatacagcatccaatgtcgaaatttaacogcgcctcagaaagcg 2653  
Qy 5533 ttcccggtgaaatagcgtcatcaaacacccctgattatcgtttatgtagaagcgaagtt 5592  
Db 2654 ttcccggtgaaatagcgtcatcaaacacccctgattatcgtttatgtagaagcgaagtt 2713  
Qy 5593 ttatgttcatatgatatatttlaactctgtgcgaatgtaaacatcagaatatttgaaaca 5652  
Db 2714 ttatgttcatatgatatatttlaactctgtgcgaatgtaaacatcagaatatttgaaaca 2773  
Qy 5653 caagctgcttcccccccccccatatttgaaagcatatacagggctatgtctcataga 5712  
Db 2774 caagctgcttcccccccccccatatttgaaagcatatacagggctatgtctcataga 2833  
Qy 5713 gcggatataatttgatgtatttagaanaataaanaataaggggttcgcgcacattc 5772  
Db 2834 gcggatataatttgatgtatttagaanaataaanaataaggggttcgcgcacattc 2893



QY 5773 cccgaaagtgccacgtgacgtctgaagaacattatcatcgtacatlaaataaa 5832  
 Db 2894 cccgaaagtgccacgtgacgtctgaagaacattatcatcgtacatlaaataaa 2953  
 QY 5833 ataggcgtatcacgagagcccttcgtc 5859  
 Db 2954 ataggcgtatcacgagagcccttcgtc 2980  
 RESULT 8  
 AAC6259 standard; DNA: 5108 BP.  
 AAC6259;  
 28-FEB-2001 (first entry)  
 Plasmid GHRH1-29YTCMV.  
 GHRH; growth hormone-releasing hormone; enzyme degradation; ds.  
 Synthetic.  
 EP1052286-A2.  
 15-NOV-2000.  
 12-APR-2000; 2000EP-0302790.  
 12-APR-1999; 99US-0128830.  
 (Pfizer ) Pfizer Prod Inc.  
 Morsey MA, Sheppard MG;  
 WPI: 2001-026585/04.  
 New polypeptide variants of growth hormone releasing hormone with enhanced resistance to enzymatic degradation, useful for treating growth hormone deficiency related disorders or to improve growth and performance -  
 Examples; Page 41-43; 67pp: English.  
 The present invention relates to growth hormone-releasing hormone (GHRH) variants having enhanced resistance to enzymatic degradation. The variant GHRH polypeptides can be administered to animals to treat growth hormone deficiency related disorders, or to improve growth and/or performance. The variants can be included in pharmaceutical compositions to promote expression and elevation of higher levels than prior art variants modified using traditional chemical methods. They have enhanced resistance to enzymatic degradation, therefore have increased length of activity.  
 Sequence 5108 BP; 1279 A; 1313 C; 1230 G; 1286 T; 0 other;  
 Query Match 36.7%; Score 2149.4; DB 22; Length 5108;  
 Best Local Similarity 94.5%; Pred. No. 0;  
 Matches 2293; Conservative 0; Mismatches 1; Indels 133; Gaps 1;

QY 3746 cataggtccgcccccgcgagacgacatcacaaaatcgacgctcaagtgagcgga 3805  
 Db 734 cataggtccgcccccgcgagacgacatcacaaaatcgacgctcaagtgagcgga 793  
 QY 3806 aaccgcagagactataaagatacagagcggttcccccgtgaagctccctcgagctc 3865  
 Db 794 aaccgcagagactataaagatacagagcggttcccccgtgaagctccctcgagctc 853  
 QY 3866 cctgttcgcagacctgcgcttaccgataccgttcgcttcctccctcggaagcg 3925  
 Db 854 cctgttcgcagacctgcgcttaccgataccgttcgcttcctccctcggaagcg 913  
 QY 3926 ggcgttccatagctcacgctgtaagtaattcgaattcggtgagtgcttcgctcaag 3985  
 Db 914 ggcgttccatagctcacgctgtaagtaattcgaattcggtgagtgcttcgctcaag 973  
 QY 3986 ctggcgctgtgacgagacccccgcttcagccgcgagccgcttcgcttaccgataat 4045  
 Db 974 ctggcgctgtgacgagacccccgcttcagccgcgagccgcttcgcttaccgataat 1033  
 QY 4046 cgtctgagtcacacccggtgaagacacgactatcgccactgagcagccactgtgaac 4105  
 Db 1034 cgtctgagtcacacccggtgaagacacgactatcgccactgagcagccactgtgaac 1093  
 QY 4106 aggatagcagagcgaggtatgtgagcggtgtgacagagttcttgaagtgtgtgctaac 4165  
 Db 1094 aggatagcagagcgaggtatgtgagcggtgtgacagagttcttgaagtgtgtgctaac 1153  
 QY 4166 tacggtacacataaagagacgatttgtatctgcttcgcttcgtaagcgaattaccctc 4225  
 Db 1154 tacggtacacataaagagacgatttgtatctgcttcgcttcgtaagcgaattaccctc 1213  
 QY 4226 ggaagaaagttgtgtaactcttgatccggaacaaacacccgctgtgagcggtgttt 4285  
 Db 1214 ggaagaaagttgtgtaactcttgatccggaacaaacacccgctgtgagcggtgttt 1273  
 QY 4286 ttgtttcagcagcagatatacgcgcagaaaagatctcaagaagatcccttgatc 4345  
 Db 1274 ttgtttcagcagcagatatacgcgcagaaaagatctcaagaagatcccttgatc 1333  
 QY 4346 ttcttcaagggtctgcagctcagtggaagcaaacacgcttaagagattgtgtcatg 4405  
 Db 1334 ttcttcaagggtctgcagctcagtggaagcaaacacgcttaagagattgtgtcatg 1393  
 QY 4406 agattatcaaaaagatcttaccctagatccctt----- 4440  
 Db 1394 agattatcaaaaagatcttaccctagatccctt----- 4440  
 QY 4441 ----- 4441  
 Db 1514 cctatctcagcagatctgtatcttctgcatcatcagttgcttgaactcggtgagggg 1573  
 QY 4453 ccgtgaggtctgctcgtggaagagtggttgctgactatcacagcgctgaatcgccca 4512  
 Db 1574 ccgtgaggtctgctcgtggaagagtggttgctgactatcacagcgctgaatcgccca 1633  
 QY 4513 tcatccagcagaagatgagggagcagcggtgtgatgagagcttctgtgtggtgagcag 4572  
 Db 1634 tcatccagcagaagatgagggagcagcggtgtgatgagagcttctgtgtggtgagcag 1693  
 QY 4573 ttgtgattttgaactttgtcttcgacggaagagtgctgtgtcggaagatgagtg 4632  
 Db 1694 ttgtgattttgaactttgtcttcgacggaagagtgctgtgtcggaagatgagtg 1753  
 QY 4633 atctgatctcttaactcagcaaaagtgctgatttatcaaaaagcgcgtcccgtaag 4692  
 Db 1754 atctgatctcttaactcagcaaaagtgctgatttatcaaaaagcgcgtcccgtaag 1813



QY 4693 tcacggtatgctctgcagctggttaccaaccaattaccattctgattagaanaacatcat 4752  
|||||  
Db 1814 tcacggtatgctctgcagctggttaccaaccaattaccattctgattagaanaacatcat 1873  
QY 4753 cgaagcatcaaatgaaactgcaattatcatatcaagaattatcaattaccattttgaa 4812  
|||||  
Db 1874 cgagcatcaaatgaaactgcaattatcatatcaagaattatcaattaccattttgaa 1933  
QY 4813 aaagcgtttctgttaagtaagagaanaacctacagaggaatttcattgagtggaagat 4872  
|||||  
Db 1934 aaagcgtttctgttaagtaagagaanaacctacagaggaatttcattgagtggaagat 1993  
QY 4873 cctgtatcgtctgcagcttccgactcgtccaaacatcaatacaactatattccctc 4932  
|||||  
Db 1994 cctgtatcgtctgcagcttccgactcgtccaaacatcaatacaactatattccctc 2053  
QY 4933 cgtcaaaaaataaggtatcaagaatgagaataccaatgagtgagcagctgataccggttga 4992  
|||||  
Db 2054 cgtcaaaaaataaggtatcaagaatgagaataccaatgagtgagcagctgataccggttga 2113  
QY 4993 atggcaaaaggttatgcaattcttccagactgtgttcaacaggccagcattacgctcgt 5052  
|||||  
Db 2114 atggcaaaaggttatgcaattcttccagactgtgttcaacaggccagcattacgctcgt 2173  
QY 5053 catcaaaatcactcgcataccaaccaaccgttatctatctcgtgattgcgctgcagcagac 5112  
|||||  
Db 2174 catcaaaatcactcgcataccaaccaaccgttatctatctcgtgattgcgctgcagcagac 2233  
QY 5113 gaaatacgcgactcgtgtttaaaggaacaattacaacaggaatcgaaatgcgaacggcgca 5172  
|||||  
Db 2234 gaaatacgcgactcgtgtttaaaggaacaattacaacaggaatcgaaatgcgaacggcgca 2293  
QY 5173 ggaacacgcgcagcgcataccaacaattttacacctgaatcagatattcttctaact 5232  
|||||  
Db 2294 ggaacacgcgcagcgcataccaacaattttacacctgaatcagatattcttctaact 2353  
QY 5233 ggaatgcgtttcccgaggagtcgcagtggtgagtaaccaatgacatcacaagagcagca 5292  
|||||  
Db 2354 ggaatgcgtttcccgaggagtcgcagtggtgagtaaccaatgacatcacaagagcagca 2413  
QY 5293 taaaatgctgattgctcggaagagagcacaattccgtcagccagtttgcgtgaccatct 5352  
|||||  
Db 2414 taaaatgctgattgctcggaagagagcacaattccgtcagccagtttgcgtgaccatct 2473  
QY 5353 catcgtatacatcattggtcaacgctactctgtgcattgttttggaaacaaactctgcagat 5412  
|||||  
Db 2474 catcgtatacatcattggtcaacgctactctgtgcattgttttggaaacaaactctgcagat 2533  
QY 5413 cgggctcccatcacaatcagatagatgtgcgacccgtgattgcccgcacatlatcgcgagccc 5472  
|||||  
Db 2534 cgggctcccatcacaatcagatagatgtgcgacccgtgattgcccgcacatlatcgcgagccc 2593  
QY 5473 atttataccataataatcagcatcccatggttggaatttaattcggcgctcgcgacaagac 5532  
|||||  
Db 2594 atttataccataataatcagcatcccatggttggaatttaattcggcgctcgcgacaagac 2653  
QY 5533 ttccctgtgaaatgagtcacatacaacccctgttattactgttttgaagagacagtt 5592  
|||||  
Db 2654 ttccctgtgaaatgagtcacatacaacccctgttattactgttttgaagagacagtt 2713  
QY 5593 ttattgtcatgatatatattttatctgtgcaatgtaacatcagagattttgagaa 5652  
|||||  
Db 2714 ttattgtcatgatatatattttatctgtgcaatgtaacatcagagattttgagaa 2773  
QY 5653 caagctggtctcccccceccatattgaagcatcattcaagggttatgtctcatgaa 5712  
|||||  
Db 2774 caagctggtctcccccceccatattgaagcatcattcaagggttatgtctcatgaa 2833  
QY 5713 gggggtacatttgaatttgaagaaataaacaataaggggttcgcgcacaatttc 5772  
|||||  
Db 2834 gggggtacatttgaatttgaagaaataaacaataaggggttcgcgcacaatttc 2893  
QY 5773 ccggaagaagtgcaccctgacgtctaagaaaccaattatlatcagacattaacctataaa 5832

Db 2894 ccggaagaagtgcaccctgacgtcttaagaaaccattattatcatgacattacaataa 2953  
|||||  
QY 5833 ataggctatcaacagagcccttcgtc 5859  
|||||  
Db 2954 ataggctatcaacagagcccttcgtc 2980  
|||||  
RESULT 9  
AAC86262  
ID AAC86262 standard; DNA; 5111 BP.  
XX  
XX AAC86262;  
AC  
XX  
XX 28-FEB-2001 (first entry)  
DT  
XX  
XX CMV IE promoter enhancer elements.  
DE  
XX  
XX GHRH; growth hormone-releasing hormone; enzyme degradation; ds.  
KW  
XX  
XX Synthetic.  
OS  
XX  
XX EPI052286-A2.  
PN  
XX  
XX 15-NOV-2000.  
PD  
XX  
XX 12-APR-2000; 2000EP-0302790.  
PF  
XX  
XX 12-APR-1999; 99US-0128830.  
PR  
XX  
XX (PFI2 ) PFIZER PROD INC.  
PA  
XX  
XX Morsey MA, Sheppard MG;  
PI  
XX  
XX WPI: 2001-026585/04.  
DR  
XX  
XX  
XX New polypeptide variants of growth hormone releasing hormone with  
PT growth resistance to enzymatic degradation, useful for treating  
PT growth hormone deficiency related disorders or to improve growth and  
PT performance  
PT  
XX  
XX Examples; Page 47-49; 67pp; English.  
PS  
XX  
XX The present invention relates to growth hormone-releasing hormone  
CC (GHRH) variants having enhanced resistance to enzymatic degradation.  
CC The variant GHRH polypeptides can be administered to animals to  
CC treat growth hormone deficiency related disorders, or to improve  
CC growth and/or performance. The variants can be included in  
CC pharmaceutical compositions to promote expression and elevation of  
CC growth hormone. The variants can be produced recombinantly at much  
CC higher levels than prior art variants modified using traditional  
CC chemical methods. They have enhanced resistance to enzymatic  
CC degradation, therefore have increased length of activity.  
CC  
XX  
XX  
SQ Sequence 5111 BP; 1280 A; 1314 C; 1230 G; 1287 T; 0 other;  
Query Match 36.7%; Score 2149.4; DB 22; Length 5111;  
Best Local Similarity 94.5%; Pred. No. 0;  
Matches 2293; Conservative 0; Mismatches 1; Indels 133; Gaps 1;



QY 3746 cataagctccgccccctgaagcagcaacaacaaatcgacgctcaagtgagtgcgca 3805  
|||||  
Db 734 cataagctccgccccctgaagcagcaacaacaaatcgacgctcaagtgagtgcgca 793  
QY 3806 aacccgaagagccttaagatacagcagcgcttcccccctgaagctccctcgagctct 3865  
|||||  
Db 794 aacccgaagagccttaagatacagcagcgcttcccccctgaagctccctcgagctct 853  
QY 3866 cctgttcacagccctgcgcttaacggatacctgtccgcttctcccttcgggaagcg 3925  
Db 854 cctgttcacagccctgcgcttaacggatacctgtccgcttctcccttcgggaagcg 913  
QY 3926 ggcctttctatagctcagcgcgtgaagatacctcaagtctcggtgaagtgctgcctcaag 3965  
|||||  
Db 914 ggcctttctatagctcagcgcgtgaagatacctcaagtctcggtgaagtgctgcctcaag 973  
QY 3986 ctgggctgtgtgacgaaccccccttcagcccgacgctgcgcttctcccttcggaactat 4045  
Db 974 ctgggctgtgtgacgaaccccccttcagcccgacgctgcgcttctcccttcggaactat 1033  
QY 4046 cgtcttgaatccaaaccccgtaagacagacttaacgacactgagcagcagcagtgtaac 4105  
|||||  
Db 1034 cgtcttgaatccaaaccccgtaagacagacttaacgacactgagcagcagcagtgtaac 1093  
QY 4106 aggatagcagagcagagtgatgtatagcggtgctacagagttcttgaagtgcgctcaac 4165  
|||||  
Db 1094 aggatagcagagcagagtgatgtatagcggtgctacagagttcttgaagtgcgctcaac 1153  
QY 4166 taagctacacatagaagacagatattgtatctgcgtctgtctgaagccagttaacctc 4225  
Db 1154 taagctacacatagaagacagatattgtatctgcgtctgtctgaagccagttaacctc 1213  
QY 4226 ggaacaaagatgtgtagctctctgtatccggcaacaaaccccgctgaagtggtttt 4285  
Db 1214 ggaacaaagatgtgtagctctctgtatccggcaacaaaccccgctgaagtggtttt 1273  
QY 4286 ttgtttcgaagcagcagatatacgcgcgaacaaagagatccaaagagctcccttgatc 4345  
Db 1274 ttgtttcgaagcagcagatatacgcgcgaacaaagagatccaaagagctcccttgatc 1333  
QY 4346 tttcttaaggggtctgaagctcagtgtagaagaaacacacacgcttaagagatttggtatg 4405  
Db 1334 tttcttaaggggtctgaagctcagtgtagaagaaacacacacgcttaagagatttggtatg 1393  
QY 4406 agattcaacaaagagatcttcacatagatctt----- 4440  
Db 1394 agattcaacaaagagatcttcacatagatctttaaataaataagagtttaataa 1453  
QY 4441 ----- 4440  
Db 1454 atctaagatataatagtaaaactggtctgacagttacaaatgcttaacagtgagca 1513  
QY 4441 -----ggggggggggggg 4452  
Db 1514 cctatctcagcagatctgctatcttcgttcatccatagtgctgaagcgggggggggggg 1573  
QY 4453 cgcctgaaggtctgctcgtgaagaaggtgtgtgactactaacaaagcctgaaatcgccca 4512  
Db 1574 cgcctgaaggtctgctcgtgaagaaggtgtgtgactactaacaaagcctgaaatcgccca 1633  
QY 4513 tcatccagcagaagaagtgaaggagcacaaggttgatagagccttctgtgaagtgagcaag 4572  
Db 1634 tcatccagcagaagaagtgaaggagcacaaggttgatagagccttctgtgaagtgagcaag 1693  
QY 4573 ttgtgtatttgaaactttgtcttgccagcgaacggtctgcgtgtgcgggaagatgctg 4632  
Db 1694 ttgtgtatttgaaactttgtcttgccagcgaacggtctgcgtgtgcgggaagatgctg 1753  
QY 4633 atctgatccttcaactcaagcaaaagtgctgattatcacaacaaagcgcgctccgcaag 4692  
Db 1754 atctgatccttcaactcaagcaaaagtgctgattatcacaacaaagcgcgctccgcaag 1813  
QY 4693 tcaaggtatgtctcgcagagtgtaacaaacatlaacaaatctgtatlaagaaactcat 4752

Db 1814 tcaaggtatgtctcgcagtgtaacaaacatlaacaaatctgtatlaagaaactcat 1873  
QY 4753 cgaagatcaaaatgaactgcaatttcatcatcagagatataacaaatattttgaa 4812  
Db 1874 cgaagatcaaaatgaactgcaatttcatcatcagagatataacaaatattttgaa 1933  
QY 4813 aaagccgcttctgtatgaagagagaaacccaacagagcagcttccaatagatgcaagat 4872  
Db 1934 aaagccgcttctgtatgaagagagaaacccaacagagcagcttccaatagatgcaagat 1993  
QY 4873 cctgtatcgtgtcgtgaattccgaactcgtccacaatcaataacaaacttaatttccct 4932  
Db 1994 cctgtatcgtgtcgtgaattccgaactcgtccacaatcaataacaaacttaatttccct 2053  
QY 4933 cgtcaaaaaataggtatcaagtgagaaatcaacatgagtagcagctgaaatccggtgaga 4992  
Db 2054 cgtcaaaaaataggtatcaagtgagaaatcaacatgagtagcagctgaaatccggtgaga 2113  
QY 4993 atgcaaaaagcttatgcatcttctccagactgtttcaacagcgcaacttaacgctcgt 2173  
Db 2114 atgcaaaaagcttatgcatcttctccagactgtttcaacagcgcaacttaacgctcgt 2233  
QY 5053 catcaaatcatcctcgatcaacaaacgcttaattcatcgtatgtgcgctgagcgagac 5112  
Db 2174 catcaaatcatcctcgatcaacaaacgcttaattcatcgtatgtgcgctgagcgagac 2233  
QY 5113 gaaatacgcagtcgctgttaaaagacaaattcaaacaggagatcgaaatcgacgcgca 5172  
Db 2234 gaaatacgcagtcgctgttaaaagacaaattcaaacaggagatcgaaatcgacgcgca 2293  
QY 5173 ggaacactgacagcgacatcaacaataatttcaacgtgaatcagaatattcttcaact 2353  
Db 2294 ggaacactgacagcgacatcaacaataatttcaacgtgaatcagaatattcttcaact 2353  
QY 5293 ggaatcgttttccgggagatcgacgtgtgtgagtaacatgacatcaacagagtagcga 5292  
Db 2354 ggaatcgttttccgggagatcgacgtgtgtgagtaacatgacatcaacagagtagcga 2413  
QY 5293 taaatgcttatgtctgagagagacataaatccgctgaagcaggtttagctgacacatc 5352  
Db 2414 taaatgcttatgtctgagagagacataaatccgctgaagcaggtttagctgacacatc 2473  
QY 5353 catctgtaacatcatgagcaagctacacttgcacgtttcagaacaaactctgagcat 5412  
Db 2474 catctgtaacatcatgagcaagctacacttgcacgtttcagaacaaactctgagcat 2533  
QY 5413 cgggtcccatcaacatgagatagatgtgcgaactgagtcgcaacatatacgcgagccc 5472  
Db 2534 cgggtcccatcaacatgagatagatgtgcgaactgagtcgcaacatatacgcgagccc 2593  
QY 5473 attataccatataaatcagacatcagatgttggaatttaacgcgcgctcgaagcagc 5532  
Db 2594 attataccatataaatcagacatcagatgttggaatttaacgcgcgctcgaagcagc 2653  
QY 5533 ttcccggtgaatgagctcacaacaaacccctgtatattacggtttatgtgaagcaggaagtt 5592  
Db 2654 ttcccggtgaatgagctcacaacaaacccctgtatattacggtttatgtgaagcaggaagtt 2713  
QY 5593 ttatgtttatgatatataattttatctgtgcaatglaacatcaagatatttgagaca 5652  
Db 2714 ttatgtttatgatatataattttatctgtgcaatglaacatcaagatatttgagaca 2773  
QY 5653 caacgctgcttccccccccccatattgaagcaattatcaaggtattgtctatga 5712  
Db 2774 caacgctgcttccccccccccatattgaagcaattatcaaggtattgtctatga 2833  
QY 5713 gcggtatcatattgagatttatattagaanaataaacaataggggttccgcgcaattcc 5772  
Db 2834 gcggtatcatattgagatttatattagaanaataaacaataggggttccgcgcaattcc 2893  
QY 5773 ccccaaaagtgcaactgagcgtctaaagaaacatattatcatgtgcatlaaactataaa 5832  
|||||



Db 2894 cccgaaaglygccactgacgtlctaagaacattatcatatgacattaaataa 2953  
Qy 5833 ataggctatcacgaggcccttcgtc 5859  
|||||  
Db 2954 ataggctatcacgaggcccttcgtc 2980  
  
RESULT 10  
AAC86254  
ID AAC86254 standard; DNA; 5185 BP.  
XX  
AC AAC86254;  
XX  
DT 28-FEB-2001 (first entry)  
XX  
DE pGHRH-4 construct.  
XX  
KM GHRH; growth hormone-releasing hormone; enzyme degradation; ds.  
XX  
OS Synthetic.  
XX  
PN EPI052286-A2.  
XX  
PD 15-NOV-2000.  
XX  
PF 12-APR-2000; 2000EP-0302790.  
XX  
PR 12-APR-1999; 99US-0128830.  
XX  
PA (PFIZ ) PFIZER PROD INC.  
XX  
PI Morsey MA, Sheppard MG;  
XX  
DR WPI; 2001-026585/04.  
XX  
XX New polypeptide variants of growth hormone releasing hormone with  
PT enhanced resistance to enzymatic degradation, useful for treating  
PT growth hormone deficiency related disorders or to improve growth and  
PT performance -  
XX  
XX Examples: Page 32-34; 67pp; English.  
XX  
CC The present invention relates to growth hormone-releasing hormone  
CC (GHRH) variants having enhanced resistance to enzymatic degradation.  
CC The variant GHRH polypeptides can be administered to animals to  
CC treat growth hormone deficiency related disorders, or to improve  
CC growth and/or performance. The variants can be included in  
CC pharmaceutical compositions to promote expression and elevation of  
CC growth hormone. The variants can be produced recombinantly at much  
CC higher levels than prior art variants modified using traditional  
CC chemical methods. They have enhanced resistance to enzymatic  
CC degradation, therefore have increased length of activity.  
XX  
SQ Sequence 5185 BP; 1302 A; 1326 C; 1260 G; 1297 T; 0 other;  
  
Query Match 36.7%; Score 2149.4; DB 22; Length 5185;  
Best Local Similarity 94.5%; Pred. No. 0;  
Matches 2293; Conservative 0; Mismatches 1; Indels 133; Gaps 1;

Db 734 |||||  
cataggtcgcgcccccgacgagcatcacaaaatcgacgtcgaagtcagagtgcgca 793  
Qy 3806 aaccgcgaagactataaagatacgaagcgtttcccccctggaagtcctctgtagcct 3865  
|||||  
Db 794 aaccgcgaagactataaagatacgaagcgtttcccccctggaagtcctctgtagcct 853  
|||||  
Qy 3866 cctgtccgaacctgccccttaccggatacctgtctccgcttctcccttcggaaagcgtg 3925  
|||||  
Db 854 cctgtccgaacctgccccttaccggatacctgtccgcttctcccttcggaaagcgtg 913  
|||||  
Qy 3926 ggcgtttctatagctcaagctgtatagttatctcagttcgtgtgaagtcgttcgctcaag 3985  
|||||  
Db 914 ggcgtttctatagctcaagctgtatagttatctcagttcgtgtgaagtcgttcgctcaag 973  
|||||  
Qy 3986 ctggagctgtgtgcaagaacccccgttgaagccgagcgtgctgcttccgtagactat 4045  
|||||  
Db 974 ctggagctgtgtgcaagaacccccgttgaagccgagcgtgctgcttccgtagactat 1033  
|||||  
Qy 4046 cgtcttgagtcacaacccgtaagaacagactatccgcaactgtgcagcagccactgtatac 4105  
|||||  
Db 1034 cgtcttgagtcacaacccgtaagaacagactatccgcaactgtgcagcagccactgtatac 1093  
|||||  
Qy 4106 aggaattagcagagcgaggtatgtagtcggtgtacagaagttcttgaagtggtgacctaac 4165  
|||||  
Db 1094 aggaattagcagagcgaggtatgtagtcggtgtacagaagttcttgaagtggtgacctaac 1153  
|||||  
Qy 4166 taaggtctactagaagaacagatttgttatctgctgctctgtctgaagccagttactcttc 4225  
|||||  
Db 1154 taaggtctactagaagaacagatttgttatctgctgctctgtctgaagccagttactcttc 1213  
|||||  
Qy 4226 ggaataaagattgtgtagctctgtatccgcgcaacaacacacccgctgtgtagcggtgttt 4285  
|||||  
Db 1214 ggaataaagattgtgtagctctgtatccgcgcaacaacacacccgctgtgtagcggtgttt 1273  
|||||  
Qy 4286 ttgtttgcaagcagcagattacgcgcgagaaaaaaggattcgaagaagtccttgatc 4345  
|||||  
Db 1274 ttgtttgcaagcagcagattacgcgcgagaaaaaaggattcgaagaagtccttgatc 1333  
|||||  
Qy 4346 ttcttaaggggtcgtgaagctcagtggaacgaatacagttgaaggtttgttcagt 4405  
|||||  
Db 1334 ttcttaaggggtcgtgaagctcagtggaacgaatacagttgaaggtttgttcagt 1393  
|||||  
Qy 4406 agattatcaaaaaggattcttaccagatccttt----- 4440  
|||||  
Db 1394 agattatcaaaaaggattcttaccagatccttttaataatgaagttaataca 1453  
|||||  
Qy 4441 ----- 4440  
|||||  
Db 1454 atctaagatatatagtaaaactgtctgcagcttaccaaagtcttaacagtgtagcgca 1513  
|||||  
Qy 4441 -----gggggggggggg 4452  
|||||  
Db 1514 cctatctcagcgatctgtctatcttcgtlccatccatagttgctcgactcggggggggggg 1573  
|||||  
Qy 4453 cgttgaggtctgcctcgtgtaagaaggtgtgtgtgactcatcacagcgccgaatccgccca 4512  
|||||  
Db 1574 cgttgaggtctgcctcgtgtaagaaggtgtgtgtgactcatcacagcgccgaatccgccca 1633  
|||||  
Qy 4513 tcatccagccagaaggtgaaggagccacggtgtgatgaaggtctgttgaagtgacaag 4572  
|||||  
Db 1634 tcatccagccagaaggtgaaggagccacggtgtgatgaaggtctgttgaagtgacaag 1693  
|||||  
Qy 4573 ttgtgtatttgaactttgtcttgcacaggaagaggtcgtggtgtgaggaagatgagcgtg 4632  
|||||  
Db 1694 ttgtgtatttgaactttgtcttgcacaggaagaggtcgtggtgtgaggaagatgagcgtg 1753  
|||||  
Qy 4633 atctgaccttcaactcagcaaaagtctcgattatccaacaagaacgcccgtccgtccaag 4692  
|||||  
Db 1754 atctgaccttcaactcagcaaaagtctcgattatccaacaagaacgcccgtccgtccaag 1813  
|||||  
Qy 4693 tcaagcgttaatgtctgcagagtttataaccaaatcaaatctctattagaanaaatcat 4752  
|||||



QY	5833	ataggcgatcaagagcccttcg	5859
Db	2954	ataggcgatcaagagcccttcg	2980

XX  
AC AAC86266;

XX Plasmid GHRH1-44YWTMV.  
DEXX  
OS Synthetic.

XX  
PN EP1052286-A2

XX 15-NOV-2000.  
PD

XX  
PF 12-APR-2000; 2000EP-0302790.

XX  
PR 12-APR-1999; 99US-0128830.

XX  
PA (PFIZ.) PFIZER PROD INC.

XX Morsey MA, Sheppard MG;  
PI

XX  
DR WPI; 2001-026585/04

XX New polypeptide variants of growth hormone releasing hormone with PT  $\alpha$ -amino acid degradation useful for treating

PT enhanced resistance to enzymatic degradation, whereas non-PT growth hormone deficiency related disorders or to improve growth and performance -

XX Examples; Page 54-56; 67pp; English  
PS

XX The present invention relates to growth hormone-releasing hormone (GHRH) variants having enhanced resistance to enzymatic degradation. CC The variant GHRH polypeptides can be administered to animals to CC treat growth hormone deficiency related disorders, or to improve CC growth and/or performance. The variants can be included in CC pharmaceutical compositions to promote expression and elevation of CC growth hormone. The variants can be produced recombinantly at much CC higher levels than prior art variants modified using traditional CC chemical methods. They have enhanced resistance to enzymatic CC degradation, therefore have increased length of activity.

Sequence 5188 BP; 1303 A; 1326 C; 1260 G; 1299 T; 0 other;

Query Match 36.7%; Score 2149.4; DB 22; Length 5188;

Best Local Similarity	94.58;	Pred. No. 0;
Matches 2293; Conservative	0;	Mismatches 1; Indels 133; Gaps 1.

3566 gcttcctcagtcactgactcgcctgcctcggtcgttcgcgtgcgcgagcgcgtatcagct 3

Db 554 gcttcctcgctaactgaactgcgtcgctcggttcgttcgctgcgcgcgcgcgtatcagct 613

3626 cactcaaaagcggtatacaggttatccacgaatcagggatacgcgaggaagacatg 3685

Db 614 cactcaaaqgcgtaatacgtttaccacgaatcagggataacgcaggaagaacatg 673

3686 tgaacaaaagccagcaaaaagccaggaaccgtaaaaaagccgcgcttgcgtgctgttctc 3745

Db 674 tgaacaaaagccacgcaaaaagccaggaaccgtaaaaagccgcgtgtcgtgcgtttttc 733

3746 cataauctccgcccccctgaagcagcatcacaaaatcgaagctcaagtcagagtgcgca 3805



Db 734 cataggctccgcccccttgacgagca tcaaaaaatcgacgtctcaagtcagaggtgcgca 793  
 QY 3806 aaccgcagagactataaagatatcagagcgcttcccccttgaaagctccccctgagcgtctc 3865  
 Db 794 aaccgcagagactataaagatatcagagcgcttcccccttgaaagctccccctgagcgtctc 833  
 QY 3866 cctgttccgacccctgcgcttaccgagatacctgttccgcttctcccttccttcgggaaagcgtg 3925  
 Db 854 cctgttccgacccctgcgcttaccgagatacctgttccgcttctcccttccttcgggaaagcgtg 913  
 QY 3926 ggccttctcctaagctcgaacgctgtatgtatctcaagttcgggtgtaggttcgttcgccaag 3985  
 Db 914 ggccttctcctaagctcgaacgctgtatgtatctcaagttcgggtgtaggttcgttcgccaag 973  
 QY 3986 ctggagctgtgtgacgaaccccccttcagccccgagacgcgcgccttatccggttaacctat 4045  
 Db 974 ctggagctgtgtgacgaaccccccttcagccccgagacgcgcgccttatccggttaacctat 1033  
 QY 4046 cgtcttgagttccaaacccggtlaaagacacagactlactgcacacttgagcaagccaactgtgtaac 4105  
 Db 1034 cgtcttgagttccaaacccggtlaaagacacagactlactgcacacttgagcaagccaactgtgtaac 1093  
 QY 4106 aggaattagcagagcgaagctatgtatgagcgtgtgtacagagttcttgaaagtgtggtcctaac 4165  
 Db 1094 aggaattagcagagcgaagctatgtatgagcgtgtgtacagagttcttgaaagtgtggtcctaac 1153  
 QY 4166 taaggtctacactagaagagcaagatctgtgtatctgcgtctgtctgttgagccaagttaaccttc 4225  
 Db 1154 taaggtctacactagaagagcaagatctgtgtatctgcgtctgtctgttgagccaagttaaccttc 1213  
 QY 4226 ggaaaaaagagttgtgtagctctctgtatccggtcaaaacacacccgcttgagcaggttggtttt 4285  
 Db 1214 ggaaaaaagagttgtgtagctctctgtatccggtcaaaacacacccgcttgagcaggttggtttt 1273  
 QY 4286 ttgtgttgcagcagcagatctacgtgcgcagaaaaaaggatctcaagaagatcccttgatc 4345  
 Db 1274 ttgtgttgcagcagcagatctacgtgcgcagaaaaaaggatctcaagaagatcccttgatc 1333  
 QY 4346 ttcttcaaggggtctgaacctcagttggaacgaataactcaagtttaaggatcttgggtcatg 4405  
 Db 1334 ttcttcaaggggtctgaacctcagttggaacgaataactcaagtttaaggatcttgggtcatg 1393  
 QY 4406 agattatcaaaaaagatctctcaactagatccctt ----- 4440  
 Db 1394 agattatcaaaaaagatctctcaactagatccctttaataaataagaagttaataatca 1453  
 QY 4441 ----- 4440  
 Db 1454 atctaaagatatatagtaaaccttgctgcaggttacccaatgcttaacagtgaagca 1513  
 QY 4441 -----ggggggggggggg 4452  
 Db 1514 cctaatctcagcgatcgtctattcgttca tccaatgctgcgtgaactcggggggggggggg 1573  
 QY 4453 cgtctgaggtctgcgtcgttgaagaaggctgtgtcgtgactcatcacagagccttgaatcgcgccca 4512  
 Db 1574 cgtctgaggtctgcgtcgttgaagaaggctgtgtcgtgactcatcacagagccttgaatcgcgccca 1633  
 QY 4513 tcatccaagcagaagaatgtgagggagccaaggtgataagagaactctgtgtatgagtgagacag 4572  
 Db 1634 tcatccaagcagaagaatgtgagggagccaaggtgataagagaactctgtgtatgagtgagacag 1693  
 QY 4573 ttggtatatttgaaactttgcttgcgaacggaacggtctcgtctgtctgcggaaagatggtg 4632  
 Db 1694 ttggtatatttgaaactttgcttgcgaacggaacggtctcgtctgtctgcggaaagatggtg 1753  
 QY 4633 atctgtatccttcaactcgaacaaagtctgaattatccaacaaagccgcgttcccgtaag 4692  
 Db 1754 atctgtatccttcaactcgaacaaagtctgaattatccaacaaagccgcgttcccgtaag 1813  
 QY 4693 tcaaggtatagctctgcagtggttataacaaatcaacaaatctctgtatgaagaaactcat 4752  
 Db 1814 tcaaggtatagctctgcagtggttataacaaatcaacaaatctctgtatgaagaaactcat 1873

QY 4753 cgaagcatcaaatgaaaactgcaatttcatlcatcagagattatcaataacatatttga 4812  
 Db 1874 cgaagcatcaaatgaaaactgcaatttcatlcatcagagattatcaataacatatttga 1933  
 QY 4813 aaagccgttctgttlaatgaagagaaactcacgcagagcgacttccataggaatggaagat 4872  
 Db 1934 aaagccgttctgttlaatgaagagaaactcacgcagagcgacttccataggaatggaagat 1993  
 QY 4873 cctgtatcgtctgttgatctccgactcgttccaaatcaataacatatttccct 4932  
 Db 1994 cctgtatcgtctgttgatctccgactcgttccaaatcaataacatatttccct 2053  
 QY 4933 cgtcaaaaataaaggttatacaagtgaagaatcacatatagtagagacggaatccggtgaga 4992  
 Db 2054 cgtcaaaaataaaggttatacaagtgaagaatcacatatagtagagacggaatccggtgaga 2113  
 QY 4993 atggcaaaaagcttatgcatcttcttccagactgtgttcaaaagcgagcca ttaacgtcgt 5052  
 Db 2114 atggcaaaaagcttatgcatcttcttccagactgtgttcaaaagcgagcca ttaacgtcgt 2173  
 QY 5053 catcaaaatcactcgcataccaacacggtatctatctgtatgttgccctgagcgagac 5112  
 Db 2174 catcaaaatcactcgcataccaacacggtatctatctgtatgttgccctgagcgagac 2233  
 QY 5113 gaataacgcgtcgttlaaaaggaattacaacaaaggaatcgaaatgcgaacggcgca 5172  
 Db 2234 gaataacgcgtcgttlaaaaggaattacaacaaaggaatcgaaatgcgaacggcgca 2293  
 QY 5173 ggaacactgcagcgca tccaacataatttcaacctgaaatcaggaatatcttctaatacct 5232  
 Db 2294 ggaacactgcagcgca tccaacataatttcaacctgaaatcaggaatatcttctaatacct 2353  
 QY 5233 ggaatcgtgttcccggggagatcgcagtggttgagtaacctgtcatctcagggagtaagga 5292  
 Db 2354 ggaatcgtgttcccggggagatcgcagtggttgagtaacctgtcatctcagggagtaagga 2413  
 QY 5293 taaatgtcttgatggtctcggaagagga taaatccgcgcagccaggtttagctgaacatc 5352  
 Db 2414 taaatgtcttgatggtctcggaagagga taaatccgcgcagccaggtttagctgaacatc 2473  
 QY 5353 catctgtaacatcatttggcaacgctaaccttctgcgatgttttcaagaaacaaactctgycgat 5412  
 Db 2474 catctgtaacatcatttggcaacgctaaccttctgcgatgttttcaagaaacaaactctgycgat 2533  
 QY 5413 cgggcttcccatcaacacggtatgattgtgcgaacctgtgtcccgacattatcgcgagccc 5472  
 Db 2534 cgggcttcccatcaacacggtatgattgtgcgaacctgtgtcccgacattatcgcgagccc 2593  
 QY 5473 attatataccataataaactcagcatcga tgttgaatttaacgcgcgcctcgagcaagcg 5532  
 Db 2594 attatataccataataaactcagcatcga tgttgaatttaacgcgcgcctcgagcaagcg 2633  
 QY 5533 ttcccggtgaataatggtcatalaacaacccctgtatctatgttbatgtatgtaagcagacgtt 5592  
 Db 2654 ttcccggtgaataatggtcatalaacaacccctgtatctatgttbatgtatgtaagcagacgtt 2713  
 QY 5593 ttattgttaatgataatttttatctgtgtgaatgtaacctcaagaattttgagaca 5652  
 Db 2714 ttattgttaatgataatttttatctgtgtgaatgtaacctcaagaattttgagaca 2773  
 QY 5653 caacgtgacttcccccccccccatatttgaagcaatttcaaggttatgttctcgtga 5712  
 Db 2774 caacgtgacttcccccccccccatatttgaagcaatttcaaggttatgttctcgtga 2833  
 QY 5713 ggcgatacatattgaaatgtatgtatga aaaaataaacaataggggtctcgcgcacatctc 5772  
 Db 2834 ggcgatacatattgaaatgtatgtatga aaaaataaacaataggggtctcgcgcacatctc 2893  
 QY 5773 cccgaaagtgcaacccggaagctcgaagaacccattatcatcagacattcaacctataaa 5832  
 Db 2894 cccgaaagtgcaacccggaagctcgaagaacccattatcatcagacattcaacctataaa 2953







QY	4753	cgagcatcaatagaactcgataatttatctatatacggatattcaattacatttttgaa	4812
Db	1874	cgagcatcaatagaactcgataatttatctatatacggatattcaattacatttttgaa	1933
QY	4813	aaagcgttctcttaatgaagggaactccacggaggaattccataggtgtaagat	4872
Db	1934	aaagcgttctcttaatgaagggaactccacggaggaattccataggtgtaagat	1993
QY	4873	ccctgtaatcgtctgcgattccgactcgtctcaacaatcaatacctaattattccct	4932
Db	1994	ccctgtaatcgtctgcgattccgactcgtctcaacaatcaatacctaattattccct	2053
QY	4933	cgtcaaaaataaggttatcaagtgaagaaatccacatgaatgaagactccggatga	4992
Db	2054	cgtcaaaaataaggttatcaagtgaagaaatccacatgaatgaagactccggatga	2113
QY	4993	atggcaaaagcttaigattcttcttcagaactgttcaacaggccagcatlaagctgt	5052
Db	2114	atggcaaaagcttaigattcttcttcagaactgttcaacaggccagcatlaagctgt	2173
QY	5053	catcaaaatcacgcgtcaatacacaacacgttattcatctgattgcgcctgaacgac	5112
Db	2214	catcaaaatcacgcgtcaatacacaacacgttattcatctgattgcgcctgaacgac	2233
QY	5113	gaatatcgatcgctgtttaaaggacaaattcaacaacaggaatcgaaatgcgaacg	5172
Db	2234	gaatatcgatcgctgtttaaaggacaaattcaacaacaggaatcgaaatgcgaacg	2293
QY	5173	ggaaacatcgacgagcatcaacaatatcttcaacctgaatcaggaattcttctaact	5232
Db	2294	ggaaacatcgacgagcatcaacaatatcttcaacctgaatcaggaattcttctaact	2353
QY	5233	ggaatgtgtcttcccgaggaatcgagtggttaagtaacatgacatcatcaggaatagca	5292
Db	2354	ggaatgtgtcttcccgaggaatcgagtggttaagtaacatgacatcatcaggaatagca	2413
QY	5293	taaaagtcttgatggtctcgagaaggagcaataatccgtcacagcaattgattgtgacact	5352
Db	2414	taaaagtcttgatggtctcgagaaggagcaataatccgtcacagcaattgattgtgacact	2473
QY	5353	catctggaacatattggcaagctctccctttgcatggtttgaagaaacaaccttgagcat	5412
Db	2474	catctggaacatattggcaagctctccctttgcatggtttgaagaaacaaccttgagcat	2533
QY	5413	cgagcttcccatcaaatcgaatagatgctgcacactgattgcccgcacattatcgcgagcc	5472
Db	2534	cgagcttcccatcaaatcgaatagatgctgcacactgattgcccgcacattatcgcgagcc	2593
QY	5473	attatacccatataaatacgaatccatgattggaatttaatcgcgcgactcgagaagaagc	5532
Db	2594	attatacccatataaatacgaatccatgattggaatttaatcgcgcgactcgagaagaagc	2653
QY	5533	tttccgttggaatataggtctataaacaccttgataactggtttatgtaagcagaagtc	5592
Db	2654	tttccgttggaatataggtctataaacaccttgataactggtttatgtaagcagaagtc	2713
QY	5593	ttattgttcatagataataatttttctctgagcaatgaaacatacagaagatttggagca	5652
Db	2714	ttattgttcatagataataatttttctctgagcaatgaaacatacagaagatttggagca	2773
QY	5653	caacgtgacttcccccccccccatattatgtagcaattatatacaggttatgtctcatga	5712
Db	2774	caacgtgacttcccccccccccatattatgtagcaattatatacaggttatgtctcatga	2833
QY	5713	ggagatcatatttgaaatgtatattagaaaaaataacaataaggggttccgcgcacatttc	5772
Db	2834	ggagatcatatttgaaatgtatattagaaaaaataacaataaggggttccgcgcacatttc	2893
QY	5773	ccgaaaaagtgccacactcgagctcctaagaaacattttatctatgcatctaaacctataaa	5832
Db	2894	ccgaaaaagtgccacactcgagctcctaagaaacattttatctatgcatctaaacctataaa	2933
QY	5833	ataggcgtatcacaggcccttctgcctc	5859

Db 2954 atagcgctatcacgagagcccttcgtc 2980

## RESULT 13

AAAF30340  
ID AAAF30340 standard; cDNA; 4779 BP.

AC AAF30340;

DT 14-MAY-2001 (first entry)

DE Expression plasmid VR1051.

KW Flt-3 ligand; Fms-like tyrosine kinase; mouse; vaccine;

KW VR1051; pBlank; rabbit; vector; ss.

OS Chimeric - Cytomegalovirus.

XX

XX

XX

XX

XX

XX

XX

XX

PT polynucleotide and one or more antigen, or cytokine encoding

PT autoimmune diseases (e.g. rheumatoid arthritis) -

PS Disclosure; Page 111-114; 149pp; English.

CC The present sequence is that of expression plasmid VR1051 (pBlank

CC enhancer, and 5' untranslated sequences followed by a multiple

transcriptional terminator region includes the polyA and

CC is used as a DNA mass control for VR6200 (see AAF30313), which

tyrosine kinase (Flt-3 ligand). The invention is directed to

cytokine by administering in vivo a Flt-3 ligand-encoding

cytokine-encoding polynucleotides.

SQ Sequence 4779 BP; 1219 A; 1203 C; 1139 G; 1218 T; 0 other;

Quorum Watch  
36 69  
C0000 C1AC B  
CD CD  
Research 1770

Best Local Similarity 94.48; Pred. No. 0;  
Matches 300; Correctly 0; Mismatches 300; Total 300

C:\Users\joseph\AppData\Local\Microsoft\Windows\Temporary Internet Files\Content.IE5\F6H8V7LW

[illegible]

On 3676 cactcaaaaccccccttatacccccacaaaatcacccccctaacccccccccccccctc 3685

Dh 0168 caactcaaaacccatataaccattatcacaacaaaatcacacccaataaacctta 227

Ov 3686 taaacaaaatcccaaaaaaacctataaatcccccccatttc 3745

\_\_\_\_\_



```
Db 2228 tgagcaaaagccagcaaaagccaggaacgcttaaaagccgctgttgctgcttttc 2287
QY 3746 catagctccgccccctgacagca tcaaaaaatcgacgtcaagttagagtggcga 3805
Db 2288 cataagctccgccccctgacagca tcaaaaaatcgacgtcaagttagagtggcga 2347
QY 3806 aaccgacagactataaagataaccggggttccccgggaagctccctcgctcct 3865
Db 2348 aaccgacagactataaagataaccgggttccccgggaagctccctcgctcct 2407
QY 3866 cctgttcggaacctgcgcttaacggatccgttccgcttcccttcggaagcggtg 3925
Db 2408 cctgttcggaacctgcgcttaacggatccgttccgcttcccttcggaagcggtg 2467
QY 3926 ggcctttccataagctacagctgtgagttatcctcgtgtgaggtcgttcgcaag 3985
Db 2468 ggcctttccataagctacagctgtgagttatcctcgtgtgaggtcgttcgcaag 2527
QY 3986 ctggagctgtgacagaaaccccccttcagcccgacccgtgcgcttaccgtaactat 4045
Db 2528 ctggagctgtgacagaaaccccccttcagcccgacccgtgcgcttaccgtaactat 2587
QY 4046 cgtcttgaagtcacaacccggttaagacagactatccgcaactgacagacgaactgtaac 4105
Db 2588 cgtcttgaagtcacaacccggttaagacagactatccgcaactgacagacgaactgtaac 2647
QY 4106 aggtatgaagagcgaggtatgtagtcggtacagagttctgtgaagtggtgcctaac 4165
Db 2648 aggtatgaagagcgaggtatgtagtcggtacagagttctgtgaagtggtgcctaac 2707
QY 4166 taaggtacataagagaagacgtattgtgtatctgtcgttcgcaacgacttaacttc 4225
Db 2708 taaggtacataagagaagacgtattgtgtatctgtcgttcgcaacgacttaacttc 2767
QY 4226 ggaaaaaaggttgtagctgtcgttcgcaacaaacacccgctgtgtagcggtttt 4285
Db 2768 ggaaaaaaggttgtagctgtcgttcgcaacaaacacccgctgtgtagcggtttt 2827
QY 4286 ttgtttgcaagcagacgtatccgcgcaaaaaaagatctcagaagatctctgtac 4345
Db 2828 ttgtttgcaagcagacgtatccgcgcaaaaaaagatctcagaagatctctgtac 2887
QY 4346 ttcttcaggggtctgacgtcagtggaacgaaactcgcgttlaaggaatttgcgtacg 4405
Db 2888 ttcttcaggggtctgacgtcagtggaacgaaactcgcgttlaaggaatttgcgtacg 2947
QY 4406 agattacaaaaagatcttaccctagacctt----- 4440
Db 2948 agattacaaaaagatcttaccctagaccttlaaataaatacgaagtttaacta 3007
QY 4441 ----- 4440
Db 3008 atctaaagtataatgagtaactgtgtcgcaggttaaccaatgcttaacgtgagca 3067
QY 4441 -----gaggggggggggg 4451
Db 3068 cctatctcagcgtatcgtatcttgcgtacatccatagtgctgcgactccggggggggggg 3127
QY 4452 ggcgttcgaggtctgcgttcggaagaagtggttcgactcaaacagcgttaacgccc 4511
Db 3128 ggcgttcgaggtctgcgttcggaagaagtggttcgactcaaacagcgttaacgccc 3187
QY 4512 atcatccagcagaagtgagggagccaggttgatgaagactgtgtgtgtagtgagca 4571
Db 3188 atcatccagcagaagtgagggagccaggttgatgaagactgtgtgtgtagtgagca 3247
QY 4572 gttgtgtatttgaactttgtcttgcaaggaacggtcgtcgttcggaagaatgctgt 4631
Db 3248 gttgtgtatttgaactttgtcttgcaaggaacggtcgtcgttcggaagaatgctgt 3307
QY 4632 gatctgatcctcaactcagcaaaagtgcattatcaacaagaacggtcccgctcaa 4691
Db 3308 gatctgatcctcaactcagcaaaagtgcattatcaacaagaacggtcccgctcaa 3367

QY 4692 gtccagtgatgctctgcaggtgttaacaaccaattaaccaattcgtattagaanaactca 4751
Db 3368 gtccagtgatgctctgcaggtgttaacaaccaattaaccaattcgtattagaanaactca 3427
QY 4752 tcggacatcaaatgaactcgtcaatttcatatcaagattatcaatcaatatttga 4811
Db 3428 tcggacatcaaatgaactcgtcaatttcatatcaagattatcaatcaatatttga 3487
QY 4812 aaaagcgtttctgttaatgaggaaggaanaactcaccgggggaggttcgaatagatggcaga 4871
Db 3488 aaaagcgtttctgttaatgaggaaggaanaactcaccgggggaggttcgaatagatggcaga 3547
QY 4872 tctgtgttcggttcggtcgtatccgactcgtccaaatcaatcaatcaatatttccc 4931
Db 3548 tctgtgttcggttcggtcgtatccgactcgtccaaatcaatcaatcaatatttccc 3607
QY 4932 tcgtcaaaaataagttatcaagtgagaatccatcgaatgagagactcgtgtgag 4991
Db 3608 tcgtcaaaaataagttatcaagtgagaatccatcgaatgagagactcgtgtgag 3667
QY 4992 aatggcaaaagttatgcaattccttcagactgttcaagaagccagcatlaacgtcg 5051
Db 3668 aatggcaaaagttatgcaattccttcagactgttcaagaagccagcatlaacgtcg 3727
QY 5052 tcatcaaaatcactcgcacatcaacaaacggttatcattcgtgtggtcgtgcgagca 5111
Db 3728 tcatcaaaatcactcgcacatcaacaaacggttatcattcgtgtggtcgtgcgagca 3787
QY 5112 cgaataacgagatcgcgtgttaaaaggaacaattacaacaggaatcgaaatgcacggcg 5171
Db 3788 cgaataacgagatcgcgtgttaaaaggaacaattacaacaggaatcgaaatgcacggcg 3847
QY 5172 aggaacactgcgcagcatcaacaatattccactcgtatccggtatccgttataacc 5231
Db 3848 aggaacactgcgcagcatcaacaatattccactcgtatccggtatccgttataacc 3907
QY 5232 tggaaatcgttttcccgagatccgcagtggtgagtaaacatcatcagagtagcg 5291
Db 3908 tggaaatcgttttcccgagatccgcagtggtgagtaaacatcatcagagtagcg 3967
QY 5292 ataaatgcttgatgctcggaagaggaataatccgtcagcgaagtttgctgacacac 5351
Db 3968 ataaatgcttgatgctcggaagaggaataatccgtcagcgaagtttgctgacacac 4027
QY 5352 tcaatcgttaaatatattgggaacgctacacttgcaggttccagaaacaactctgcgca 5411
Db 4028 tcaatcgttaaatatattgggaacgctacacttgcaggttccagaaacaactctgcgca 4087
QY 5412 tcgggtcccatatacatcgaatgagtgctgcacgtatgcccgaacatcatatccggagcc 5471
Db 4088 tcgggtcccatatacatcgaatgagtgctgcacgtatgcccgaacatcatatccggagcc 4147
QY 5472 catatacccatataaatacagcatccagtgtggaatttaacgcggccttcgagaagac 5531
Db 4148 catatacccatataaatacagcatccagtgtggaatttaacgcggccttcgagaagac 4207
QY 5532 gttcccggttgataatggtcctaaacacccctgtatatacgtttatgtagagcaaggt 5591
Db 4208 gttcccggttgataatggtcctaaacacccctgtatatacgtttatgtagagcaaggt 4267
QY 5592 ttatattgtcatgagatataatttatctgtgtgaacgttaacatcaagaatttggagac 5651
Db 4268 ttatattgtcatgagatataatttatctgtgtgaacgtttatgtagagac 4327
QY 5652 acaagctgcttcccccccccccatatgtgaacattatcaagggttatgtctcatg 5711
Db 4328 acaagctgcttcccccccccccatatgtgaacattatcaagggttatgtctcatg 4387
QY 5712 agcgaatacatattgaatgtatttgaanaaataaagaatgggttcgcgcacatt 5771
Db 4388 agcgaatacatattgaatgtatttgaanaaataaagaatgggttcgcgcacatt 4447
```







OY	4632	gacatgctccctcaactcagcaaaagtctgatttatttcaacaagccgcctcccgtaa	4691
Db	3635	gactcgtacctcaactcagcaaaagtctgatttatttcaacaagccgcctcccgtaa	3694
OY	4692	gtcagcgtaatgctcttcgcaggtttcaacaacattacaacattctgattagaanaacta	4751
Db	3695	gtcagcgtaagccttcgcaggtttcaacaacattacaacattctgattagaanaacta	3754
OY	4752	tcgagcatcaaatgaaactgcgaatttattcatatcaggtattatcattaccatttttga	4811
Db	3755	tcgagcatcaaatgaaactgcgaatttattcatatcaggtattatcattaccatttttga	3814
OY	4812	aaaagccgttctctgtaatgaaggagaaacacacccgggcaggttccataagatgcaaga	4871
Db	3815	aaaagccgttctctgtaatgaaggagaaacacacccgggcaggttccataagatgcaaga	3874
OY	4872	tccttggttcggtctgcgagcttcgcactctgcacaacatcaaaccttaatttcccc	4931
Db	3875	tccttggttcggtctgcgagcttcgcactctgcacaacatcaaaccttaatttcccc	3934
OY	4932	tcgctcaaaaataagttatcaagtgagaaatcacacatgagtgagacgttgatccggtgag	4991
Db	3935	tcgctcaaaaataagttatcaagtgagaaatcacacatgagtgagacgttgatccggtgag	3994
OY	4992	aatggcaaaaagcttatgacattcttcttccagaacttglttaaaagggccagccattacgctcg	5051
Db	3995	aatggcaaaaagcttatgacattcttcttccagaacttglttaaaagggccagccattacgctcg	4054
OY	5052	tcatcaaaaactactgcgcatcaacacacccgttatcttcatcttggtgccttgagcaga	5111
Db	4055	tcatcaaaaactactgcgcatcaacacacccgttatcttcatcttggtgccttgagcaga	4114
OY	5112	cgaatacgcgacgcgcgtgttaaaagagacaatttacaacacaggaatcgaatgcacccgcgc	5171
Db	4115	cgaatacgcgacgcgcgtgttaaaagagacaatttacaacacaggaatcgaatgcacccgcgc	4174
OY	5172	aggacaactgcgacgcgcatcaacaatattttaactcggtaactcaggtatcttcttaatcc	5231
Db	4175	aggacaactcgcgcgcgcatcaacaatattttaactcggtaactcaggtatcttcttaatcc	4234
OY	5232	tggaaatcgtggtttcccgggagtcgcagtggttgagtaaacatgcatcactcagagtgtaagg	5291
Db	4235	tggaaatcgtggtttcccgggagtcgcagtggttgagtaaacatgcatcactcagagtgtaagg	4294
OY	5292	ataaaatgcttgatggtctcggaagggaggaataattccgctcgacggagtttgctgacacac	5351
Db	4295	ataaaatgcttgatggtctcggaagggaggaataattccgctcgacggagtttgctgacacac	4354
OY	5352	tcaatctgtaaatcaatttgcaacgcttaaccttgcatggttttcagaaacaacctctggcga	5411
Db	4355	tcaatctgtaaatcaatttgcaacgcttaaccttgcatggttttcagaaacaacctctggcga	4414
OY	5412	tcgggcttcccatatacatcgatagattgtgcacacctgattggccggaattatctgcagcc	5471
Db	4415	tcgggcttcccatatacatcgatagattgtgcacacctgattggccggaattatctgcagcc	4474
OY	5472	catattacccaataaatacagacatccatgtgtggaatttaactcgcgcgctcgcagcaagc	5531
Db	4475	catattacccaataaatacagacatccatgtgtggaatttaactcgcgcgctcgcagcaagc	4534
OY	5532	gtttcccggttgaatatggtctcaaacocccctgtatactacggtttatgtaagcagacagt	5591
Db	4535	gtttcccggttgaatatggtctcaaacocccctgtatactacggtttatgtaagcagacagt	4594
OY	5592	tttaatgttcaatagatagatatattttatctctgcaatgtaaacatcagagatttgagac	5651
Db	4595	tttaatgttcaatagatagatatattttatctctgcaatgtaaacatcagagatttgagac	4654
OY	5652	acaacggtgcttcccccoccccccacatttgaagcatttaccaggtttatgttctcatg	5711
Db	4655	acaacggtgcttcccccoccccccacatttgaagcatttaccaggtttatgttctcatg	4714
OY	5712	agcggaatacatatttgaatgattttagaanaataaacaataaggggttccgcgcacattc	5771



Dh 2557 tgagcaaaagccagcaaaagccagaaacgtlaaaaggccgctgtcgtgcytttttc 2616  
Qy 3746 cataggtccgccccctgaaagagatcacaaaaatcgagctcaagtcagaagcgagcga 3805  
Dh 2617 cataggtccgccccctgaaagagatcacaaaaatcgagctcaagtcagaagcgagcga 2676  
Qy 3806 aaccgagagactataaagatcacagcgctttcccccgtgaagctccctctgtgacct 3865  
Dh 2677 aaccgagagactataaagatcacagcgctttcccccgtgaagctccctctgtgacct 2736  
Qy 3866 ccgtgtccgaccctgacgtcttaaccggatacgtgtccgacctctccctcgggaaagctg 3925  
Dh 2737 ccgtgtccgaccctgacgtcttaaccggatacgtgtccgacctctccctcgggaaagctg 2796  
Qy 3926 ggcgtttctcatagctcagctcgtgtatgtatctcaagttccgtgttaagtcgttcgcccaa 3985  
Dh 2797 ggcgtttctcatagctcagctcgtgtatgtatctcaagttccgtgttaagtcgttcgcccaa 2856  
Qy 3986 ctgggctgtgtgcagaaacccccgttcagcccgacccgctgcgaccttaaccgtaactat 4045  
Dh 2857 ctgggctgtgtgcagaaacccccgttcagcccgacccgctgcgaccttaaccgtaactat 2916  
Qy 4046 cgtcttgagttccaacccggttaagacagactatcgcacatgtagcagccacatgtgtaac 4105  
Dh 2917 cgtcttgagttccaacccggttaagacagactatcgcacatgtagcagccacatgtgtaac 2976  
Qy 4106 aggatatacagaagcgagtgatgttagcgcggtgtctaaagagttcttaagtgtgagcctaac 4165  
Dh 2977 aggatatacagaagcgagtgatgttagcgcggtgtctaaagagttcttaagtgtgagcctaac 3036  
Qy 4166 taagctcactaagaagaagacagatttgttatctgcgctcgtctgaagccagttacccttc 4225  
Dh 3037 taagctcactaagaagaagacagatttgttatctgcgctcgtctgaagccagttacccttc 3096  
Qy 4226 ggaaaaagagttgtagctcttgatccgcgcaaaacacacccgctgtgtagcggtgtttt 4285  
Dh 3097 ggaaaaagagttgtagctcttgatccgcgcaaaacacacccgctgtgtagcggtgtttt 3156  
Qy 4286 ttgtgttcagaagcagcaattacgacgcaaaaaaagatctccaagaagatctctgtatc 4345  
Dh 3157 ttgtgttcagaagcagcaattacgacgcaaaaaaagatctccaagaagatctctgtatc 3216  
Qy 4346 ttcttaacggggtcgtgaagctcgaatggaacgaaacacacgtttaagggaatttgcgtcat 4405  
Dh 3217 ttcttaacggggtcgtgaagctcgaatggaacgaaacacacgtttaagggaatttgcgtcat 3276  
Qy 4406 agattatacaaaaagactctcaactagatcctt----- 4440  
Dh 3277 agattatacaaaaagactctcaactagatccttlttaaaatgaagttttaatca 3336  
Qy 4441 ----- 4440  
Dh 3337 atctaagtatatatgagtaaaactgtgtcagcagttaccaaigtcttaatcagtgagcga 3396  
Qy 4441 ----- 4451  
Dh 3397 cctatcctcagcagctgtctatcttcgtcatcatcagatgttcgtgactccggggggggggg 3456  
Qy 4452 ggcgtgagagctgtcgtcgttgaagaagtggtgtgcatactacacagagcctgaacgcgcc 4511  
Dh 3457 ggcgtgagagctgtcgtcgttgaagaagtggtgtgcatactacacagagcctgaacgcgcc 3516  
Qy 4512 atcatccagccagaagaatggaaggagccagcggttgaatgaagcttggttgaagtgcagca 4571  
Dh 3517 atcatccagccagaagaatggaaggagccagcggttgaatgaagcttggttgaagtgcagca 3576  
Qy 4572 gttgtgtgattttgaacttttgcgttgccacaggaagcggtcgtgtgcgggaagatgcgt 4631  
Dh 3577 gttgtgtgattttgaacttttgcgttgccacaggaagcggtcgtgtgcgggaagatgcgt 3636  
Qy 4632 gatcgtatcccttaactaagcaaaaagttcogatttatccaagaacgcgcgtcccgctcaa 4691  
Dh 3637 gatcgtatcccttaactaagcaaaaagttcogatttatccaagaacgcgcgtcccgctcaa 3696

Qy 4692 gtcacgttaatgtctcgtccaggtttacaaaccaattaaaccaattctgtatagaanaactca 4751  
Dh 3697 gtcacgtttaaigtctcgtccaggtttacaaaccaattaaaccaattctgtatagaanaactca 3756  
Qy 4752 tcgaagcatcaaattgaaactgcgaatttatcatatacagaatlaacatacatattttga 4811  
Dh 3757 tcgaagcatcaaattgaaactgcgaatttatcatatacagaagttatcaatacatattttga 3816  
Qy 4812 aaaagccgtttctgttaatgaaggagaaacccaacgagagcagttccaatagatgycaga 4871  
Dh 3817 aaaagccgtttctgttaatgaaggagaaacccaacgagagcagttccaatagatgycaga 3876  
Qy 4872 tccgtgtatcgggttcggagttccgactcgttccaacaatcaatacaacttaattttccc 4931  
Dh 3877 tccgtgtlatcgggttcggagttccgactcgttccaacaatcaatacaacttaattttccc 3936  
Qy 4932 tcgtcaaaaataaagttatacgaagtgaagaatcaccaatgagtgaagatgaatccggtga 4991  
Dh 3937 tcgtcaaaaataaagttatacgaagtgaagaatcaccaatgagtgaagatgaatccggtga 3996  
Qy 4992 aatggaacaaagcttatgtcatcttccgaagcttgttcaacagcgccacctaagctcg 5051  
Dh 3997 aatggaacaaagcttatgtcatcttccgaagcttgttcaacagcgccacctaagctcg 4056  
Qy 5052 tcataaaaatcactcgcacccaacaaacggttatcttcattcgtgattgcgcctgagcagga 5111  
Dh 4057 tcataaaaatcactcgcacccaacaaacggttatcttcattcgtgattgcgcctgagcagga 4116  
Qy 5112 cgaataacgcgaatcgcgtgttaaaagagacaattatacaaacaggaatcgaacgcgcgc 5171  
Dh 4117 cgaataacgcgaatcgcgtgttaaaagagacaattatacaaacaggaatcgaacgcgcgc 4176  
Qy 5172 aggaacacgtccagcgcatcaacaataattccaactgaaatcaggaatatcttccataacc 5231  
Dh 4177 aggaacacgtccagcgcatcaacaataattccaactgaaatcaggaatatcttccataacc 4236  
Qy 5232 tggaaatgcgttttcccggggatacgaatggtgagatgaacatgcatcacaagaagtagcg 5291  
Dh 4237 tggaaatgcgttttcccggggatacgaatggtgagatgaacatgcatcacaagaagtagcg 4296  
Qy 5292 atcaaatgtctgattgtcgcgaagagcaataatctccgtacgcaggttttagctgcacac 5351  
Dh 4297 atcaaatgtctgattgtcgcgaagagcaataatctccgtacgcaggttttagctgcacac 4356  
Qy 5352 tcaatcgttaacatcattggaacagcttaaccttgccatggtttcaagaaacatctggcgca 5411  
Dh 4357 tcaatcgttaacatcattggaacagcttaaccttgccatggtttcaagaaacatctggcgca 4416  
Qy 5412 tcgggcttcccaataacatcagatgtgtcgcacccgtgattgcccgcacatataccgagcc 5471  
Dh 4417 tcgggcttcccaataacatcagatgtgtcgcacccgtgattgcccgcacatataccgagcc 4476  
Qy 5472 catltaaccaataaatacagatccatcattgtgaaatttaacgcggccctcgagcaagac 5531  
Dh 4477 catltaaccaataaatacagatccatcattgtgaaatttaacgcggccctcgagcaagac 4536  
Qy 5532 gtttcccggttgaatatgtcgtcataaacacccctgtatactgtttatgttaagcagacgt 5591  
Dh 4537 gtttcccggttgaatatgtcgtcataaacacccctgtatactgtttatgttaagcagacgt 4596  
Qy 5592 ttattatggtcatgatatatttttatctgttgcgaatgtaacaatcagaagattttgaagc 5651  
Dh 4597 ttattatggtcatgatatatttttatctgttgcgaatgtaacaatcagaagattttgaagc 4656  
Qy 5652 acaacgtgtgcttcccccctccccaatlaatgaagatttatcagaaggttatgttctcag 5711  
Dh 4657 acaacgtgtgcttcccccctccccaatlaatgaagatttatcagaaggttatgttctcag 4716  
Qy 5712 agcggatatacatttggaatgtatttgaagaaataaataaggggttccgcgcacatt 5771  
Dh 4717 agcggatatacatttggaatgtatttgaagaaataaataaggggttccgcgcacatt 4776



OY 5772 ccccgaaagtgcacctgacgtctaaagaaccattattatcatgacattaaactataa 5831  
Db 4777 ccccgaaagtgcacctgacgtctaaagaaccattattatcatgacattaaactataa 4836  
OY 5832 aataggcgtatcacagagcccttcgtc 5859  
Db 4837 aataggcgtatcacagagcccttcgtc 4864

Search completed: August 20, 2002, 22:31:00  
Job time: 12448 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 17:52:31 ; Search time 103.58 Seconds  
(without alignments)  
13894.253 Million cell updates/sec

Title: US-09-826-206-3  
Perfect score: 5859  
Sequence: 1 tcgcgcgttcgcgtgatgcac.....taccacagagcccttcgc 5859

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2-6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2-6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2-6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2-6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2-6/ptodata/1/ina/PT05\_COMB.seq: \*  
6: /cgn2-6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2146.8	36.6	5900	US-08-663-998-1	Sequence 1, Appli
2	2146.8	36.6	5952	US-08-663-998-2	Sequence 2, Appli
3	2145.2	36.6	5676	US-08-663-998-3	Sequence 3, Appli
4	2145.2	36.6	5682	US-08-663-998-4	Sequence 4, Appli
5	1940	33.1	4296	US-08-316-950-15	Sequence 15, Appli
6	1940	33.1	4296	PCT-US95-12642-15	Sequence 15, Appli
7	1940	33.1	4352	US-08-316-950-16	Sequence 16, Appli
8	1940	33.1	4352	PCT-US95-12642-16	Sequence 16, Appli
9	1940	33.1	6047	US-08-316-950-12	Sequence 12, Appli
10	1940	33.1	6047	PCT-US95-12642-12	Sequence 12, Appli
11	1904	32.5	3993	US-08-316-950-14	Sequence 14, Appli
12	1904	32.5	3993	PCT-US95-12642-14	Sequence 14, Appli
13	1904	32.5	6044	US-08-316-950-18	Sequence 18, Appli
14	1904	32.5	6044	PCT-US95-12642-18	Sequence 18, Appli
15	1903	32.5	4120	US-07-977-630-20	Sequence 20, Appli
16	1903	32.5	4120	US-07-977-630-21	Sequence 21, Appli
17	1903	32.5	4120	US-08-316-950-11	Sequence 11, Appli
18	1903	32.5	4120	PCT-US95-12642-11	Sequence 11, Appli
19	1903	32.5	4505	US-08-316-950-13	Sequence 13, Appli
20	1903	32.5	4505	PCT-US95-12642-13	Sequence 13, Appli
21	1903	32.5	6171	US-08-316-950-17	Sequence 17, Appli
22	1903	32.5	6171	PCT-US95-12642-17	Sequence 17, Appli
23	1537.4	26.2	3987	US-09-082-649B-83	Sequence 83, Appli
24	1491	25.4	3987	US-09-082-649B-84	Sequence 84, Appli
25	1237.2	21.1	3418	US-08-944-916-12	Sequence 12, Appli
26	1237.2	21.1	4066	US-09-298-367B-3	Sequence 3, Appli
27	1237.2	21.1	5838	US-08-850-880-3	Sequence 3, Appli

C 28	1237.2	21.1	5838	2	US-08-944-916-3	Sequence 3, Appli
C 29	1237.2	21.1	5838	3	US-08-814-877-3	Sequence 3, Appli
C 30	1237.2	21.1	10317	2	US-09-058-746-1	Sequence 1, Appli
C 31	1237.2	21.1	10317	4	US-09-438-142-1	Sequence 1, Appli
C 32	1237.2	21.1	14113	3	US-09-223-134-1	Sequence 1, Appli
C 33	1237.2	21.1	14113	3	US-08-992-801-1	Sequence 1, Appli
C 34	1237.2	21.1	14113	4	US-09-223-535-1	Sequence 1, Appli
C 35	1236.8	21.1	4326	4	US-08-760-615-7	Sequence 7, Appli
C 36	1215.8	20.8	6407	1	US-07-977-630-1	Sequence 1, Appli
C 37	1215.8	20.8	6407	1	US-07-977-630-2	Sequence 1, Appli
C 38	1213.4	20.7	7641	1	US-07-924-028A-5	Sequence 5, Appli
C 39	1073	18.3	2115	1	US-08-329-681A-3	Sequence 3, Appli
C 40	1016.4	17.3	1360	4	US-09-082-649B-80	Sequence 80, Appli
C 41	984.4	16.8	1360	4	US-09-082-649B-81	Sequence 81, Appli
C 42	979.4	16.7	6407	1	US-07-977-630-1	Sequence 1, Appli
C 43	979.4	16.7	6407	1	US-07-977-630-2	Sequence 2, Appli
C 44	894.6	15.3	3356	1	US-08-246-403A-1	Sequence 1, Appli
C 45	894.6	15.3	3356	1	US-08-246-403A-4	Sequence 4, Appli

## ALIGNMENTS

```
RESULT 1
US-08-663-998-1
: Sequence 1, Application US/08663998
: Patent No. 5846946
:
: GENERAL INFORMATION:
: APPLICANT: HUEBNER, Robert C.
: APPLICANT: NORMAN, Jon A.
: APPLICANT: LIANG, Xiaowu
: APPLICANT: CARNER, Kristin R.
: APPLICANT: BARBOUR, Alan G.
: APPLICANT: LUKE, Catherine J.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: NUMBER OF SEQUENCES: 4
: TITLE OF INVENTION: ADMINISTERING BORRELIA DNA
: CORRESPONDENCE ADDRESS:
: ADDRESS: Curtis, Morris & Safford, P.C.
: STREET: 530 Fifth Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10036
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/663,998
: FILING DATE: 06-JUN-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Kowalski, Thomas J.
: REGISTRATION NUMBER: 32,147
: REFERENCE/DOCKET NUMBER: 454312-2440
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-840-3333
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5900 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
:
: US-08-663-998-1
:
: Query Match 36.6%; Score 2146.8; DB 2; Length 5900;
: Best Local Similarity 94.4%; Pred. No. 0;
: Matches 2292; Conservative 0; Mismatches 2; Indels 134; Gaps 1;
```



[illegible]

Db	4309	ATCATCCAGCCAGAAAGTGGAGGGAGCCCGGTGATGAGAGCGCTTGTTGTTAGGTGACC	4368
Oy	4572	gttggtaatttgaactttgctcttgccaaggaagctctgcgttctgcggagaatgct	4631
Db	4369	gTTGGTGTATTTGAACTTTTCTTGTGGCAAGGAGGTGCTTGTCGGGAAGATGCGT	4428
Oy	4632	gactcgactccttcaactcagcaaaagtctgatttatccaacaaagccgcgtcccgctca	4691
Db	4429	GATTCGATTCCTTCAACTCGCAAAAGTTCGATTTATTCACAAAGCCGCCGTCGTCGA	4488
Oy	4692	gtccacgcttaagtcctgcacgtgttacaaccaattaacccaattctcgattagaanaactca	4751
Db	4489	GTGAGGGTAATGCTCTGCCAGTGTACACCAATTAAACCAATTCGTATTAAGAAAACTGA	4548
Oy	4752	tcgagatcaaatgaaactgcaatttatctatcagattatcaataaccatatttcta	4811
Db	4549	TCGAGCATCAATTGAACGCAATTTATTTCAATATCAGGATTATCATATCCATTTTTTGA	4608
Oy	4812	aaaagcgtttcttgaatgaagaagaaacacacacagcgagcttccatagatgtgcaaga	4871
Db	4609	AAAAGCGTTTCTGTATTAAGAGAAAAACACCGAGGCGAGTTCATAGATGCGCAAGA	4668
Oy	4872	tcctgtlaticgtctcgagattccgcactgtccacaataataaacaacttaatttccccc	4931
Db	4669	TCCGATATGCGTCTCGATTCGCACTCGTCACATCATATACACCATTAATTTTCCC	4728
Oy	4932	tccttcaaaaataaggttatacgaatgagaatcacatcagatagtcagactgaatccggtgag	4991
Db	4729	TCGTCAAAAATAAGTTATTCAGATGAGAAATCACCATGATGACGACACTGAATCCGGTGG	4788
Oy	4992	aatggcaaaagcttatgcatcttcttccagactgtttcaacagcgccacacatccgcctcg	5051
Db	4789	AATGGCAAAAGCTTATGCAATTTCTTTCACAGCTTGTTCAACAGGCGACGCAATAGCGTGG	4848
Oy	5052	tcaaccaaatcaactcgcgcataccaaccaacogttatcattcgttattgcgctgtgagcaga	5111
Db	4849	TCATCAAAATACATCGCATCACCAAAACCGTTATTCATTCGATTTGGCGCTCGACGCGA	4908
Oy	5112	cgaatatacgcatcgtcgtttaaagaagcaatttacaacaggaatcgaatgtcaaacccgagcgc	5171
Db	4909	CGAAATAGCGCATGCTGTATTAAAGACATTTACAAACAGGAATCGAATGCAACCGGCGC	4968
Oy	5172	aggaacactgcagcagcatcaacaataatttccactggaatcaggaattctcttataacc	5231
Db	4969	AGGACACTGCGACGAGCATCAACATATTTTTCACGCTGAATCAAGATATTTCTTATATACC	5028
Oy	5232	tggaaatcgttttcccgaggatcgcaggtgtgtgaataccatcaccatcaatcaggaatgacgg	5291
Db	5029	TGGAAATGCTGTTTCCCGGGATGCGACATGSGTAGTAAACCATGCATCATCAGAGATACGG	5088
Oy	5292	atcaaatgtcttgaatgtcgaagaagcataaattccgttcagccaagttagtctgaccatcc	5351
Db	5089	ATAAATATGCTGATGATGTCGGAAGAGCGCTAAATTTCCGTCAGCCAGCTTATGTCGACATC	5148
Oy	5352	tcaatcgtataactcatttgcaacagcttactcttgcattgtttcagaanaaacctctgycoga	5411
Db	5149	TCATTCGTATACATCATATGSCAACGCTACTCTTGCATGTTTCAGAAACAACCTGTGGGCGCA	5208
Oy	5412	tcgagcttcccatatacatcgatagatgtgcgaactgtatcccgcaaatatctgcagagcc	5471
Db	5209	TCGGGCTTCCCATACAAATGATAGATTGTGCGCACTGATTCGCCGACATTTATTCGCAAGCC	5268
Oy	5472	catttataccatataataacacagcatccatggttgaaatttaactcgcgcgttcagacaagc	5531
Db	5269	CATTATATCCCATATTAATACAGCATCCAGTTGGAAATTTAATCGGGCGCTCGACCAAGAC	5328
Oy	5532	gttcccgcttgaatatgtgcataacaaccccttgattatacgttitaatgtaagcaagct	5591
Db	5329	GTTTCCCGGTGAAATATGCGTCATTAACACCCCTTGATTAAGCTTTATATGTAAGCAAGAGT	5388
Oy	5592	tttatgttataatataataattcttactcttgcgaatgtgaacatcagagattttgagac	5651



Db	5389	TTTATTGTCATGATCATATATTTTATATCTTGTCGACAGTAAACAACAGAGATTTTGGAC	5448
Qy	5652	acaacgcygcgttcccccacccccaatlgaagcatltaacaggtlcatgltccatg	5711
Db	5449	ACAACGTGGCTTCCCCCCCCCATTTTGTGAAGCATTTTACAGGTTATTGTCTATG	5508
Qy	5712	agcggaatacatattgaaatgatltaagaataaataaagaatagggttcgcgcacatlc	5771
Db	5509	AGCGGAATCATATTTGAAATCTATTTTGAAAAATTAACAATATGGGGTTCCCGGCACATTT	5568
Qy	5772	ccccgaagaatgccaacctgcctgcctaaagaacacatattatcatatgacattaacctataa	5831
Db	5569	CCCCGAAAATGGCCACTGACCTCTAAGAAACCATTTATTATCATTAACCTATTAA	5628
Qy	5832	aatagcgcatacaagagcccttcgctc	5859
Db	5629	AATAGCGGTATTCAGAGGCCCTTTCTCG	5656

```

1  RESULT      2
2  US-08-663-998-2
3  : Sequence 2, Application US/08663998
4  : Patent No. 5846946
5  :
6  : GENERAL INFORMATION:
7  : APPLICANT: HUBNER, Robert C.
8  : APPLICANT: NORMAN, Jon A.
9  : APPLICANT: LIANG, Xiaowu
10 : APPLICANT: CARRER, Kristin R.
11 : APPLICANT: BARBOUR, Alan G.
12 : APPLICANT: LUKE, Catherine J.
13 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
14 : NUMBER OF SEQUENCES: 4
15 : CORRESPONDENCE ADDRESS:
16 : ADDRESSEE: Curtis, Morris & Safford, P.C.
17 : STREET: 530 Fifth Avenue
18 : CITY: New York
19 : STATE: NY
20 : COUNTRY: USA
21 : ZIP: 10036
22 :
23 : COMPUTER READABLE FORM:
24 : MEDIUM TYPE: Floppy disk
25 : COMPUTER: IBM PC compatible
26 : OPERATING SYSTEM: PC-DOS/MS-DOS
27 : SOFTWARE: Patent Release #1.0, Version #1.30
28 : CURRENT APPLICATION DATA:
29 : APPLICATION NUMBER: US/08/663,998
30 : FILING DATE: 06-JUN-1996
31 : CLASSIFICATION: 424
32 :
33 : ATTORNEY/AGENT INFORMATION:
34 : NAME: Kowalski, Thomas J.
35 : REGISTRATION NUMBER: 32,147
36 : REFERENCE/DOCKET NUMBER: 454312-2440
37 : TELECOMMUNICATION INFORMATION:
38 : TELEPHONE: 212-840-3333
39 : INFORMATION FOR SEQ ID NO: 2:
40 : SEQUENCE CHARACTERISTICS:
41 : LENGTH: 5952 base pairs
42 : TYPE: nucleic acid
43 : STRANDEDNESS: single
44 : TOPOLOGY: linear
45 : MOLECULE TYPE: DNA (genomic)
46 :
47 : US-08-663-998-2

```

	Query Match	Similarity	Score	DB	Length
Best Local	94.4%	Pred. No. 0	2146.8	2	5952
Matches 2292	Conservative	0	Indels 2	Indels 134	Gaps 1
0y	3566	gcttcctcgctacgcgactcgctgctgctgcgcttcgagctcgagcgatcacgact	3625		
Db	3281	gcttctctgctgctacgcgactcgctgctgctgcttcgagctcgagcgatcacgact	3340		

QY	3626	cactcaaaagcgcgttaatacagttatccacaagaatcaggggataacgcaggaagaacacag	3665
Db	3341	CACCTAAAGGCGGGTAATACGGTTATCCACAGATACAGGGGATACGCAAGAAACATG	3400
QY	3666	tgaagcaaaagccagcgcaaaagggccaggaacgcgtaaaaagccgcgttgtcgtgttttc	3745
Db	3401	TGAGCAAAAAGCCAGCAAAAAGGCGAGAACCGTAAAAAAGCCGCGTGTGTCGCTTTTTC	3460
QY	3746	catagcgtccgcgcgcgccttgacgcagcatcacaaaatcgaagctcaagtcagagcgtgggga	3805
Db	3461	CATAGCGTCGCCCCCTCTGACAGACATACAAAATACGCTCAAGTACAGATGAGGTGGCA	3520
QY	3806	aaccgcagaagcatataagaataccagcgcttcccccctggagagctccctcgtgcgcct	3865
Db	3521	AACCGACAGCATATTAAGTATACAGGGCTTTCCCTTGAAAGCTCCCTGTGCGTCT	3580
QY	3866	ccctgtccgaaccctgcgcgttaaccgatacctcgtccgcttccctctgggaagcg	3925
Db	3581	CCTGTTCGACCCGCGCGCTTACCGGATACCTTCGCGCTTCTCCCTTCCGGAAGCGTG	3640
QY	3926	gcgccttctcaatagctcaagcgttaagtaacctagtcgtgttaagtcgttcgcaccaag	3985
Db	3641	GCGCTTTCTCAATGCTACCGCTAGGTATCTAGTTCCGTTAGTGGTTCGCTCCAG	3700
QY	3986	ctggcgtctgtgcagcaaccgccgtctcagcccgacgcgtgcgccttalcgcgttaactat	4045
Db	3701	CTGGCGCTGTGHCACGAACCCCGGTTACGCCGACCGGTGGCGCTTATCCGTTAATAT	3760
QY	4046	cgctcttgatccaaaccgcgttaagacacagcatatgcgcacgtggcagcagccactgttaac	4105
Db	3761	CGCTTTGAGTCAACCCCGGTAAACACAGACTTATCCCACTGGCAGCACCACCTGGTAAC	3820
QY	4106	aggatttaacagagcgaaggtatgttagggcgtgcacagagttcttgaagtggcgtctaac	4165
Db	3821	AGGATTACAGACGCGAGGATTGTAAGCGGTGCTACAGAGTTCTTGAGTGGTGGCTTAAC	3880
QY	4166	tacgcgtcacactagaagaagacagtaattgttatctgcgtctcgtgaagccagtaccctc	4225
Db	3881	TACGCGTACACTAGAAAGSACGATTTTGGTATCTGGCTGCTGTGAAGCAGATTACCTTC	3940
QY	4226	ggaaaaagatctgtagcctcttgatccgcgcaaaaaacacacgcgtgtgtagcgtgtttc	4285
Db	3941	GGA AAAAGAGTGGTACTCTTGTATCCGGAACAAACACACGCTGTAGCGGTGTTTT	4000
QY	4286	tttgtttgcaagcagcagatttaacgcgagaaaaaaggcttccaagaagaatcccttgatc	4345
Db	4001	TTTGTTTCAAGCAGCAGATTACCGCGA AAAAAGATCTCAAGAAATCTTTGTATC	4060
QY	4346	ttttctaaagggctctgacgtcagctggaacgaanaactcagcttlaagggaatttgtagctg	4405
Db	4061	TTTTCTACGGGCTGTGACGCTCAGTGAAGTGAAGAAACTCAGCTTAAAGGATTTTGGTATG	4120
QY	4406	agattatacaaaaagatcttcaecttcaactagatcccttt-----	4440
Db	4121	AGATTATCAAAAAGACTTCACTCACTAGTCCCTTTTAAATTAATAATGAAGTTTAAATCA	4180
QY	4441	-----	4440
Db	4181	ATCTTAAGTATATATATGAGTAACCTGGTCTGACAGTTACCAATGCTTAATACGTAGSCCA	4240
QY	4441	-----	4441
Db	4241	CCTATCTCAGGAGTCTGTCTAATTCGTATCCATATAGTTGGCTGACTCCGGGGGGGGGG	4300
QY	4452	gcgcgtgaagctcgtccgtgaagaagctgtctgtaatacaaccagcctgaatacgcgcc	4511
Db	4301	GCGGTGAAGTCTGCTGCTGTGAAGAAGTGTCTGTGACTATACACAGGCTCTGAATGCCCC	4360
QY	4512	atcatccagcgcaagaagtggagggaacacagctgtatgaagaagtttgtttagtggtgagca	4571
Db	4361	ATCATTCACCGCAAGAGTGAGGGAGCCACGCTGTGATGAAGAGCTTGTGTATAGTGGACCA	4420
QY	4572	gttggtagttcttgaaactttgcttttgcaaggaagcgtctgcgttgtcgggaagatgcgt	4631



Db	4421	GTTCGTCATTTTAACTTTTGGTTTGGCCAGGAAGGCTCCGTTGTCCGGAGATGGCG	4480
Qy	4632	gaticgaccccttaactcagcaaaagtctcgatttcaacaagccgcgtccgctca	4691
Db	4481	GATCGTACCTTCAACTCCAGCAAAAGTTGATTTATTCAMCAAAAGCCGCGCTCCGTCA	4540
Qy	4692	gtcagcgtaagtctcgcgaaggttcaaaccaattcaacaaattctgtattgaaaaactca	4751
Db	4541	GTCAAGCTAATAGTCTGCCAGTGTGTACAACCAATTAAACCAATTCTGATTGAAAACTCA	4600
Qy	4752	tcgagcaatcaaatgaaactgcgaatttatcatcatcaaggattatcaatacatatttga	4811
Db	4601	TCGAGCATCAATTCGAAATCGCAATTATTCATATCAGATTATTCGAATTCATATTTTGA	4660
Qy	4812	aaaaagccgtttctgtatagtgaagagaaaactcacagagcggttccatagatgtgcaag	4871
Db	4661	AAAAGCCGTTTCGTATATGAAGAGAAAATCTACCGAGGCGATTCATAGGATGGCAGA	4720
Qy	4872	tcctcgatcgcgtctcgattccgattccgactcgtccaaatcaatacaacctattatcccc	4931
Db	4721	TCCTCGTATCGGTCGTGCGATTTCGCAATCGTCCAAATCAATACAAACGATTAAATTTCCC	4780
Qy	4932	tcgtcaaaaaaaggttatacaagttagaataccaatgatagtgacgtgaatccggtgtag	4991
Db	4781	TCGTCAAAAAATAGTGTATATCAAGTGTGAAGATCACCATGATGTACACATCGAATCCGGTAG	4840
Qy	4992	aatggcaaaagctatgcattcttccttccttcgaactctgttcaagagcgcaactgaactcg	5051
Db	4841	AATGGCAAAAGCTTATATCATTTTCTTCCAGACTTGTTCACAGCGCAGCCATTACGCTCG	4900
Qy	5052	tcattcaaatcaactcgcgataccaaccaaccgattatcattcgtgattgcgcttgagaga	5111
Db	4901	TCATCAAAATATCATCTCGCATCAACCAACCGTTATTCATTCGTGATTCGCGCTGAGCAGA	4960
Qy	5112	cgaataacgcgatactgcctggtttaaaaagcaattacaacagagatcgaattgaacacggcgc	5171
Db	4961	CGAAATACGCGATCGCTGTTTAAAAAGACATTTAACAAACGGGATCGAATGTGACAGCCGGCGC	5020
Qy	5172	aggaacactgcacagcgatacaacaatatcttaacctgaaatcgaatattcttctaatacc	5231
Db	5021	AGGAACACTCGCACCGGATCAACAATATTTTCACTCGAATCAGATATTTCTTAAATACC	5080
Qy	5232	tggaaatcgtctttcccgaggatcgcagtggtgagttaacctgatcatcaaggagttacg	5291
Db	5081	TGGAAATCGCTTTTCCGGGGATTCGCAAGTGTAGTAACCAATGCATCAAGGATGACGG	5140
Qy	5292	atcaaatcctgtagtgcgcggaagagatcaaatcttcgtaacgcagtttagtctgatacacc	5351
Db	5141	ATCAAAATCGTGTATGTGTCGGAAGAGGATTAATTCGCTGACGCAAGTTATGTGTGACATC	5200
Qy	5352	tcattcgtacaatcattggcaacgcgtaccccttgcacatgtttcagaacaactctvggcga	5411
Db	5201	TCATCTGTAAACATATTTGGCAACCTCACTTTCCATGTTTCAGAAACAACCTCTGGCGCA	5260
Qy	5412	tcgggcttcacatacaatcgaatagattgtgcgaccgtatgcgcgacattatcvgaac	5471
Db	5261	TGGGCTTCCCATCAATCGATGATTTGTGCACTGATTGCCGACATTAATCGGGAGCC	5320
Qy	5472	caattataccaataaatacgcataccatgttggaaattaaatcgcgcctcogaagaag	5531
Db	5321	CATTATTAACCATTAATTAATCAGCATTCATGTTGAATTAATTCGCGCTCGACGAAAGAC	5380
Qy	5532	gtttcccggttgaataatgctcataaacccctgtatctgtttatgtgaagcaagct	5591
Db	5381	GTTTCCCGTGTGAATTTGGCTCATTAACACCCCTGTATACCTGTTTATGTAAACCAACAGT	5440
Qy	5592	tttatgtttcatgataatataatttctctgtgcgaatgtgaacatcaagactttgagac	5651
Db	5441	TTTATTGTTCAATGATGATATATTTTATCTTGTGCAATGTAAATCAGAGATTTTGGAGC	5500
Qy	5652	aaacgtggctttcccccctcccatatattgaagcaattatcaagggtattgtctcatg	5711

Db	5561	ACAACGTGGCTTCCCCCCCCCATTATTTGAGCATTTATCAGAGTTATGTCTCATG	5560
Oy	5712	agcgaatacatattgaatgtaattttagaaaaataaaacaataagggtctccgcacattc	5711
Db	5561	AGCGGATACATATTTTGAAATGTATTTGAAAAAATAACAATATGAGGGTTCCGCCACATTT	5560
Oy	5772	ccccgaaaagtgcacctgaactgaactctaagaacatatatatgatgacattaacctataa	5831
Db	5621	CCCCGAAAAGTCACCCTGCAGCTCTAAGAAACCATTTATATCATGACATTAACCTATATA	5680
Oy	5832	aataggcgatacacgaagagcccttcgtc	5859
Db	5681	AATAGGCGTATCACGAGGCCCTTTTCGTC	5708

```

1      RESULT      3
2      US-08-663-998-3
3      ; Sequence 3, Application US/08663998
4      ; Patent No. 5846946
5      ; GENERAL INFORMATION:
6      ; APPLICANT: HUEBNER, Robert C.
7      ; APPLICANT: NORMAN, Jon A.
8      ; APPLICANT: LIANG, Xiaowu
9      ; APPLICANT: CARNER, Kristin R.
10     ; APPLICANT: BARBOUR, Alan G.
11     ; APPLICANT: LUKE, Catherine J.
12     ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
13     ; TITLE OF INVENTION: ADMINISTERING BORRELIA DNA
14     ; NUMBER OF SEQUENCES: 4
15     ; CORRESPONDENCE ADDRESS:
16     ; ADDRESSEE: Curtiss, Morris & Safford, P.C.
17     ; STREET: 530 Fifth Avenue
18     ; CITY: New York
19     ; STATE: NY
20     ; COUNTRY: USA
21     ; ZIP: 10036
22     ; COMPUTER READABLE FORM:
23     ; MEDIUM TYPE: Floppy disk
24     ; COMPUTER: IBM PC compatible
25     ; OPERATING SYSTEM: PC-DOS/MS-DOS
26     ; SOFTWARE: Patentln Release #1.0, Version #1.30
27     ; CURRENT APPLICATION DATA:
28     ; APPLICATION NUMBER: US/08/663,998
29     ; FILING DATE: 06-JUN-1996
30     ; CLASSIFICATION: 424
31     ; ATTORNEY/AGENT INFORMATION:
32     ; NAME: Kowalski, Thomas J.
33     ; REGISTRATION NUMBER: 32,147
34     ; REFERENCE/DOCKET NUMBER: 444312-2440
35     ; TELECOMMUNICATION INFORMATION:
36     ; TELEPHONE: 212-840-3333
37     ; INFORMATION FOR SEQ ID NO: 3:
38     ; SEQUENCE CHARACTERISTICS:
39     ; LENGTH: 5676 base pairs
40     ; TYPE: nucleic acid
41     ; STRANDEDNESS: single
42     ; TOPOLOGY: linear
43     ; MOLECULE TYPE: DNA (genomic)
44     ; US-08-663-998-3

```

Query	Match	Similarity	Score	DB 2	Length
Best Local	94.4%	Pred. No. 0			
Matches	2291	Conservative	0	Mismatches	
				Indels	134
				Gaps	1
QY	3566	gcttctcgctacfractcgtctgcgctcgttcgctcgtcgctcgagcgatcagct	3625		
Db	3005	GCCTCTCGCTACACACACGCTGCGCTCGGCTCGGCTCGGCGACGGTATAGCT	3064		
QY	3626	cactcaaaagcgatatacagttatcacagaatcagggataacgcaggaagaacatcy	3685		
Db	3065	CACTCAAAGCGGTATACGGTTATCCACAGATCACGGGATATACCGAAGAACATNG	3124		



Oy 3686 tgaagcaaaagccagcaaaagccaggaacccgttaaaagccgctgtgcgtctttcc 3745  
|||||  
Db 3125 TGAGCAAAAGGCGACACAAAGGCCAGAAACCGTAAAAAGCGCGCTGTGCGCTTTTTC 3184  
Oy 3746 cataagctcgcgcgccttgcagcagcatcaaaaaatcgacgctcaagltcagagtgycga 3805  
Db 3185 CATAGGCTCCGCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGTGCGCA 3244  
Oy 3806 aaccgcgaagactataaagatacagcaggtttcccccggagagctccctgtgtgagctc 3865  
|||||  
Db 3245 AACCCGACAGGACTATTAAGATACCAAGGCGTTCCCGCTGGAAAGCTCCCTGCGGCTCT 3304  
Oy 3866 cctgttcacgacctgcgccttaccggaatacgtgtccgccttctccctcgcggaagcgtg 3925  
Db 3305 CCGTGTCCGACCGCTCGCTTACCGGATACGTCGCGCTTCTCCCTTCCGGAAAGCGTG 3364  
Oy 3926 ggcgtttctcatagctcagcgtgtgatactcagttcgggtgtgagtgctgtccccaag 3985  
Db 3365 GCGCTTCTCAATGCTCAAGCGGTGATGTCAGTTCGCTAGAGTGTGCTGCCCAAG 3424  
Oy 3986 ctggagctgtgtgacgaaccccccgcttaccgcccgcgtgcgcttaccggttaactat 4045  
|||||  
Db 3425 CTGGGCTGTGTGACGAACCCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACAT 3484  
Oy 4046 cgtcttgagttccaaaccggtlaagacagactatcgccactgacgacgacactgtgtaac 4105  
Db 3485 CGTCTTGAGTCCACACCCGGTAAGACACGACTTATCGCCACTGCGACGACCACTGGTAAC 3544  
Oy 4106 aggatlaagacagaglatgtgagcggtgtctacagagttcttgaagtgtgtccctaac 4165  
Db 3545 AGGATTACAGACGAGGATATAGCGGTGTACAGAGTCTTGAAGGTGCGCTAAC 3604  
Oy 4166 taagctcacactaagaagaagcatttgtagtctgtgcgtctgtgaagcaattacccttc 4225  
|||||  
Db 3605 TACGGCTACACTAGAAAGGACATATTTGGTATCTCGCTGCTGAAGCACTTACCTTC 3664  
Oy 4226 ggaaaaaagttgtgagctctgtatccgcgcaaaacaaaccgcgtgtgagcggttctt 4285  
Db 3665 GGAATAAGGTTGATAGCTTGATCCGGCAAAACAAACGCGGTGTAGCGGTGTTT 3724  
Oy 4286 ttgtcttgcagcagcagatctacgcgcgcaaaaaaagatctcaagaagaaccttctgac 4345  
Db 3725 TTTGTTTCAACACACAGATTTACGGCGAGAAAAAAGATCTCAAGAAATCTTTGATC 3784  
Oy 4346 ttcttcaggggtctgcagcctcaagtgtgaaacgaacactcaagttgaaggtttgtgac 4405  
Db 3785 TTTTCTACGGGCTGTGACGCTAGTGGACGAAACCTCACGTTAAGGATTTGGTCA 3844  
Oy 4406 agatatacaaaaagatcttccactagatccttt----- 4440  
|||||  
Db 3845 AGATTATCAAAAAGATCTTCACTAGATCTTTTAAATTGAATTTAAATCA 3904  
Oy 4441 ----- 4440  
Db 3905 ATCTAAAGTATATATAGTAAACTTGTCTGACAGTTACCAATGCTTAAATCAGTGAGCA 3964  
Oy 4441 ----- 4451  
Db 3965 CCTATCTACGCGATCTGTCTATTTGCTTATCCATATGTTGCTGACTCCGGGGGGGGG 4024  
Oy 4452 ggcgtgaggtctgcctcgtgaagaagtgltgtgactcatacgaagcctgtgaccccc 4511  
Db 4025 GGGCGAAGGTCTGCTCGTGAAGAAGTGTGCTGACTATACGAGCCTGATGCCCC 4084  
Oy 4512 atcatccagcgaagaagtgaaggagccacggtltgataagagcttgtgtgaagtgaaca 4571  
Db 4085 ATCATCCAGCCAGAAAGTGAGGAGCCACAGGTTGATGAGAGCTTGTGTAGTGAGACA 4144  
Oy 4572 gttgtgtatatttgaaactttgtcttgcacggaacggtgtgtgtgtgtgtgtgtgtgt 4631  
Db 4145 GTTGTGATTTTGAACCTTTGCTTGTGCCACGGAACGCTGTGCGGTGTGCGGAAGATGCG 4204  
Oy 4632 gatctgatacctcaactcagcaaaagttcgaattatcaacaagaacgcgctcccgctcaa 4691

Db 4205 GATTCGATCTCTCAACTGACGAAAAAGTTGATTTATTAACAAGCGCGCTGCCGTCAA 4264  
Oy 4692 gtcagcgtaaigtctctgcagtglttaaccaacttaaccaacttctgtatagaanaactca 4751  
Db 4265 GTGAGGTTAATCTCTGCGCAGTGTACAACCAATTAAACCAATTCTATAGAAAACTCA 4324  
Oy 4752 tcgaagatcaaaatgaanaacgcaattatcatatacgaagttatcaatacatttttga 4811  
Db 4325 TCGAGCATTAATTAAGAACTGCAATTTATTCATTCAGGATTTATCAATTTTGTGA 4384  
Oy 4812 aaaagcgtttctgttaatgaagagaanaacacacgagcgagttccataagatggaaga 4871  
Db 4385 AAAAGCGTTTCTGTGAATGAAGAGAAACTCACGAGGACAGTTCATAGATGACAGA 4444  
Oy 4872 tccgtgatacgtctgcgagttccgactgtgtccaaactcaatacactatttccccc 4931  
Db 4445 TCTGTGATCGGTCTGCGATTCCGACTCGTCCAACTCAATACMACCTATTATTTCC 4504  
Oy 4932 tcgtcaaaaataagttatcaagtgagaataccaagagtgagtgagctgaatccggtgag 4991  
Db 4505 TCGTCAAAAATTAAGTTATCAAGTGAGAAATCACATGAGTGACGACTGAATCCGGTGA 4564  
Oy 4992 aatgcaaaaagcttatgtcaattcttccagactgtgtcaacagggccagccattacgctcg 5051  
Db 4565 AATGGCAAAAGCTTATGATGATTTCTTCCAGACTGTGTTCAACAGGCCAGCATTAAGCTCG 4624  
Oy 5052 tcatcaaaatcactcgcataccaacaaacggttatcatctcgtatgtgcgtgtgagcga 5111  
Db 4625 TCATCAAAAATCACTCCATCAACCAACCGTTATTCATTCGTGATGTCCTGACGAGA 4684  
Oy 5112 cgaatacgcagatcgtgtgttaaaaggaacttaacaacgggaatgaatgcaacggcg 5171  
Db 4685 CGAATACCGGATCGCTGTTAAAGGACATTTCAACAGGAATGCAATGCAACCGCGC 4744  
Oy 5172 aggaacacgycagcgcatacaacaatatlttcaaccggaatcaggaatcttcttaataac 5231  
Db 4745 AGGAACACTGCCAGCGCATCAACAAATATTTTCACTGATACAGATATTTCTTAAATAC 4804  
Oy 5232 tggaaagcgtgttcccgaggatcgagtggtgtgagtaaccaatgacatcaagagtaag 5291  
Db 4805 TGGAAATGCTGTTTCCCGGGGATCGCAGTGTAGTAACCATCAATCAGAGATGACG 4864  
Oy 5292 ataaaaatgtctgtagtgcgaagaagcaataatctcgtcaagcaggtttagtcgacatc 5351  
Db 4865 ATTAATGCTTGTATGTCGGAAGAGCATTAATCCGTGACGCAAGTTTACTGTGACATC 4924  
Oy 5352 tcatctgtiaacatctgtgcaacgctacacttgcctatgttccagaaacaacttgcgcga 5411  
Db 4925 TCATCTGTAAATCATTTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTGCGCA 4984  
Oy 5412 tcgggcttcccatatacaatcgaatagatgtgtgcacccgtatgtcccgacatattcgagac 5471  
Db 4985 TCGGGCTTCCCATPACAAATGATGATTTGCGACCTGATTTGCCGACATTTATGCGAGCC 5044  
Oy 5472 catllaccatataataacagcatcagtttgaaatttaatcgcgccctcgagaagac 5531  
Db 5045 CATTTATACCATATAAATCAGCATCAGTGTGAAATTTAATGCGGCTTAGAGCAAGAC 5104  
Oy 5532 gtttcccgltgaatagttcgtatacaaacccctgtlatactgtltatgtlaagagacagt 5591  
Db 5105 GTTTCGGTTGAATATGGGTCATTAACACCCCTGTATTTACTGTTATGTAAGACAGAGT 5164  
Oy 5592 ttatatgttcaatgataatatttatacttctgtgcaatgtlaacaatcaagatlttgagac 5651  
Db 5165 TTTATTTTCATGATATATATTTATTTCTGTGCAATGTAAATCATGAGATTTTGAAC 5224  
Oy 5652 acaagtggtcttcccccaccccccataatgaagcatltaacgaaggttatgttccatg 5711  
Db 5225 ACAAGTGCTTCCCGCCCGCCCGCCATTTATGAAAGCATTTATCAAGGTTATTTCTCATG 5284  
Oy 5712 agcgaatacatltgaatgtatlttagaaaaataaacaataggggttccgcgacatlt 5771  
|||||



```

Db 5285 AGCGATACATATTTGAATGATTTTGAATAAATAAGGGTTCCGGCAGATT 5344
Qy 5772 cccgaaagtcacactgacgtcctaagaacatattatcatgacattaccataa 5831
Db 5345 CCCCCAAATGGCAGCTGACGTCTAAGAAACCATTTATTCATGACATTAACTATAA 5404
Qy 5832 aataagcgtatcaagagcccttcgtc 5859
Db 5405 AATAGCGTATCACGAGCCCTTTCGTC 5432

RESULT 4
US-08-663-998-4
; Sequence 4, Application US/08663998
; Patent No. 5846946
; GENERAL INFORMATION:
; APPLICANT: HUEBNER, Robert C.
; APPLICANT: NORMAN, Jon A.
; APPLICANT: LIANG, Xiaowu
; APPLICANT: CARNER, Kristin R.
; APPLICANT: BARBOUR, Alan G.
; APPLICANT: LUKE, Catherine J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: ADMINISTERING BORRELIA DNA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,998
; FILING DATE: 06-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalski, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2440
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-840-3333
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-663-998-4

Query Match 36.6%; Score 2145.2; DB 2; Length 5682;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2291; Conservative 0; Mismatches 3; Indels 134; Gaps 1;

```

```

Qy 3746 catagctccgccccctgacgagcatcaacaaaaatcgacgtcgaagtcagagtcgca 3805
Db 3191 CATAGCTCCGCCCCCTTGACGAGCATACAAAATACGAGGCTCAAGTCAGAGTGCGCA 3250
Qy 3806 aaccgacgaactatataaataacagagcgttcctccctggaagctcctctgtgctct 3865
Db 3251 AACCAGCAGAGCATATAAGATACGAGGCTTCCCTCGAAGCTCCCTCGGCTCT 3310
Qy 3866 cctgtccgaccctgacgtcctacgcgaatcttcgccccttcctccctcggaagcgtg 3925
Db 3311 CCGTTCGACCTCGCGCTTACCGATACCTGTGCGCCCTTCTCCCTTGGGAAACGCTG 3370
Qy 3926 ggccttctcatagctacacgtctgaggtatctcagttcggtgtagtgcgtcctcaag 3985
Db 3371 GCGCTTCTCAATGCTACGCTGATGATGATGATGATGATGATGATGATGATGATGAT 3430
Qy 3986 ctggcctgtgtgcaagaaaccccgcttcagcccgacgcgcgttcggtcctatccgtaact 4045
Db 3431 CTGGCCTGTGTGCAAGAACCCCGCTTACGCCGACGCTGCGCTTATCCGTAATAT 3490
Qy 4046 cgtcttgagtcacaccggttaagacagcactatccgcaactgcaagcaactgtaac 4105
Db 3491 CGCTTGAAGTCCACCCGCTAAGACAGACACTTATGCCAGCAGCAGCAGCAGCAGTAC 3550
Qy 4106 aggatagaagacgaggtatgtagcgtgtcgaagagttcgtgaagtgagtcctaac 4165
Db 3551 AGGATAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCGTGAAGTGGTGGCTTAC 3610
Qy 4166 taagctacactagaagagcaagatttgtagtctgcgtctgtgaagcgaactcttc 4225
Db 3611 TACGCTACACTAGAAAGACAGATTTTGTATCTCGCTCTGACCGATTAACCTTC 3670
Qy 4226 ggaataagagtgtagctctgacatccgcaaaaaccccgctgtagcgtgagtggttt 4285
Db 3671 GGAATAAGAGTGTGAGCTCTGATCCGCAAAACACCGCTGTAAGCGGTGTTT 3730
Qy 4286 ttgtttgcaagacagatgacgtacgcgcaaaaaaagatctcaagaagatccttgatc 4345
Db 3731 TTTGTTTGAAGACGACGATTAACGCCGCAAAAAAGATCTCAAGAAATCTTGTGATC 3790
Qy 4346 ttctcagcgggtctgacgtcagtcagtggaacgaactcaagtgtaggttttgatcgt 4405
Db 3791 TTTTCTACGGGGTCTGACGCTCAGTGAACGAATACTCAGTTAAGATTGGTCATG 3850
Qy 4406 agattatcaaaaagatcttcactagatccttt 4440
Db 3851 AGATTATCAAAAAGATCTTACCTAGATCTTTTAATTAATAAATGAAGTTTAATCA 3910
Qy 4441 ----- 4440
Db 3911 ATCTAAGTATATATGAGTAATCTTGCTGACAGTTACCAATGCTTAATCACTAGAGCA 3970
Qy 4441 ----- 3999999999 4451
Db 3971 CCTATCTACGAGCTGTCTATTTTCCTGATCACTACTGCTGACCTCGGGGGGGGG 4030
Qy 4452 ggcgtgaggtctgctcgtggaagaagtggtgctgactatacgaagcctgagatgccc 4511
Db 4031 GCGGTGAGGTCTGCTCGGAAGAAGGTGTGCTGACTCATATCAAGCCTGAAATCGCCC 4090
Qy 4512 atcatccagcgaagaagtgagggcagcaggtgatgagacgttgtagtgaagtcga 4571
Db 4091 ATCATCCAGCCGAAAGTGAAGGAGCCACGGTGAATGAGACCTTTGTTGTAGTGAGCA 4150
Qy 4572 gttagtgatctgaaacttgcttgccagaggaagcgtctgctgctcggaagatgct 4631
Db 4151 GTTGGTGAATTTTGAATTTGCTTGTGCAAGGAGGCTGCTGTTGCGGAAATGCGT 4210
Qy 4632 gactgaccccttaactcaagcaaaagttcgatttcttaaaaaaacgcgcgtccgtcaa 4691
Db 4211 GATCTGATCTTCACTCAAGCAAAAGTTGATTTATCAACAAACCGCGCTCCGCTAA 4270
Qy 4692 gtcaagctaatgtctgcagtggttacaaccaatcaatctgattagaataactca 4751

```























```

|||||
Db 1816 TATTAAGATACAGAGCGTTTCCCTCGAAGCTCCCTGCGCTCTCTCTGCGAGCC 1757
Qy 3879 tgcgcgttacccgataacacctgcgcgttccttccttcggaagcgltggcgcttcata 3938
Db 1756 TCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTCGGAGAGCGTGCGGCTTCTCAT 1697
Qy 3939 gctcacgctgttagtactcactgctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3998
Db 1696 GCTACCGCTGTAGGTATCTCAGTTCGCTGTAGGTGTGTGTGTGTGTGTGTGTGTGTGT 1637
Qy 3999 accgaaccccccttcagcccgacgctgcgcgttccttcacgtaactacgtlcttgagcca 4058
Db 1636 ACCGAGCCCCGTTACGCCCGACCGCTGCCCTTATCCGTTACATCTGCTGAGTCCCA 1577
Qy 4059 acccggtgaagacacgacttacgcacctgtgcagcagccactgtgtacagatgaagag 4118
Db 1576 ACCCGGTGAGACAGACTTATCGCACTGGCAGCAGCAGCTGTGTAACAGAGATTAGCAGAG 1517
Qy 4119 cgaaggtatgtagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4178
Db 1516 CGAGGTATGATGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1457
Qy 4179 gaagacagatattgtgtatctgcctgtgtgaagccagttacccctgcgaaagaggttg 4238
Db 1456 GAAGGACAGTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1397
Qy 4239 gtagctctgtatccggaacaaacacccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4298
Db 1396 GTAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1337
Qy 4299 accagattacgcgcgaaagaaagatctcaagaagatccttgcattcttctacgggt 4358
Db 1336 ACGAATTACGCGGAGAAAGAGATCTCAAGAGATCTTGTATCTTGTCTGAGGAGT 1277
Qy 4359 ctgaagctcagtggaagaaacacacacacacacacacacacacacacacacacacacac 4418
Db 1276 CTGAGCGCTCAGTGAACCTGATC----- 1252
Qy 4419 ggaactcaccctagatccttccttcggaagcggtgtgtgtgtgtgtgtgtgtgtgtgtgt 4478
Db 1253 ---GATTATTAAAGCTTTCGCAATTTCTGCAGCTGACTAGTAGAGTCTGCTCGTGAAGAGG 1197
Qy 4479 tgtgtctgactacacacacacacacacacacacacacacacacacacacacacacacacac 4538
Db 1196 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1137
Qy 4539 accggttgatbagagcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4598
Db 1136 ACGGTTGATGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1077
Qy 4599 cagggaaacggtctgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4658
Db 1076 CACGGAAGCGGTCTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
Qy 4659 tgcgttatcacaacaaacgcgcgtccgcgtcaagtcagcgttaagtcgtcgtcagtgtaac 4718
Db 1016 TCGATTATTACAAAGCGCGCTCCGTCGCAAGTCAGCGTATGTGTGTGTGTGTGTGTGTGT 957
Qy 4719 aaccaatcaacaaatctgtatgaaacacacacacacacacacacacacacacacacacac 4778
Db 956 AACCAATTAAACCAATTTCTGATTGAAAAACCTATCGAGCAATCAAAATGAAACTCAATT 897
Qy 4779 ttcataatagatataatataatataatataatataatataatataatataatataatata 4838
Db 896 TTCAATATGAGATATATCAATATCAATATTTTGAAGAACCGCTTCTGTATGAAGAGAA 837
Qy 4839 aactcacgagcgagttccatagagatgagcaagatcctgtgtatcgtgtgtgtgtgtgtgt 4898
Db 836 AACTCAGCAGGAGATCTCATAGATGTGCAAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 777
Qy 4899 cgtcaaatcaatacaacacacacacacacacacacacacacacacacacacacacacacac 4958

```

```

Db 776 CGTCCAACTAATFACACCTATTATTCCTCCCTGTCAAAATAAGTTATCAAGTAG 717
Qy 4959 aaatcacatagatgagacactgaatcccggtgagaaatggaagacgttctgactcttc 5018
Db 716 AATATCAGATGATGAGAGACTGAAATCCGGAGAAATGGCAAAAGATTATGCTTTCTTC 657
Qy 5019 cagactgttcaacagcgacacacacacacacacacacacacacacacacacacacacacac 5078
Db 656 CAGACTTGTTCACAGCGCCAGCCATTAAGCTGTCTATCAAAATCACTGCATCAACAA 597
Qy 5079 ccgttatcattctgtatgtgcctcagagagagacgaataacgcgactcgtgtgtgtgtgtgt 5138
Db 596 CCGTATTATTCATTCGTGATTCGCTGAGCAGAGAGAAATCGGATGCTGTGTAAGAG 537
Qy 5139 caatcaaacaggaatcgaatgtcaacggcgaggaacacacacacacacacacacacacacac 5198
Db 536 CAATTACAAACAGAAATGAAATGAAACCGGCGCAGAAACCTCCAGCGCATCAACAA 477
Qy 5199 tttacactgatacagatattcttctaatacctgtgaatgcgttctccggagatcgca 5258
Db 476 TTTTACCTGAAATCAGATATTCTTAAATCTGAAATGCTGTGTGTGTGTGTGTGTGTGT 417
Qy 5259 gtygtgaatcaacatgcaatcagagatgagatgagatgagatgagatgagatgagatgag 5318
Db 416 GTGTGAGTAAACATGATCATGATGAGAGTACGATGAAATGCTGTGATGCTGTGAAAGAG 357
Qy 5319 ataatttcgtagcagcagatgagatgagatgagatgagatgagatgagatgagatgagatg 5378
Db 356 ATTAATTCGATGAGCAGATTTAGTGTGACCATCTCATCTGTAATGAGTGTGAAACCT 297
Qy 5379 ccttgcacatgttcaagaacacacacacacacacacacacacacacacacacacacacac 5438
Db 296 CCTTTCCTCATGTTTCAAGAAACATGCGGATGAGGCTTCCATACACCGATGAT 237
Qy 5439 gtgcacacacacacacacacacacacacacacacacacacacacacacacacacacacac 5498
Db 236 GTGCACTGATGAGCCGCAATTAATGAGCCATTAATGAGCCATTAATGAGCCATTAAT 177
Qy 5499 atgttgaaatlaatcgcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5558
Db 176 ATGTGGAATTTAATCGCGCTCGAGCAAGACGTTTCCGTTGAAATGCTGTGATGAG 117
Qy 5559 cccctgtatctatgttataagacagacacacacacacacacacacacacacacacacac 5618
Db 116 CCCCCTGATTAATGTTATGTAAGCAGACAGTTTATGTTATGATGATGATGATGATGAT 57
Qy 5619 tcttgycaatgtaacatcagagatttgagacacacacacacacacacacacacacacac 5664
Db 56 TCTTGTCAATGTAACTAGAGATTTTGTGACACAAACGTCCTT 11

```

```

RESULT 9
US-08-316-950-12/c
; Sequence 12, Application US/08316950
; Patent No. 5679515
GENERAL INFORMATION:
APPLICANT: STOVER, Charles K.
TITLE OF INVENTION: Mycobacterial Reporter Strains and Uses
TITLE OF INVENTION: thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Khourie and Crew
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```



```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/316,950
3 FILING DATE: 03-OCT-1994
4 CLASSIFICATION: 435
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Hunter, Tom
7 REGISTRATION NUMBER: 38,498
8 REFERENCE/DOCKET NUMBER: 15371A-001100
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (415) 543-9600
11 TELEFAX: (415) 543-5043
12 INFORMATION FOR SEQ ID NO: 12:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 6047 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: DNA (plasmid)
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: 1..6047
22 OTHER INFORMATION: /standard_name="plasmid_pmh30-lux"
23
24 US-08-316-950-12

```

Query Match	33.18;	Score 1940;	DB 1;	Length 6047;
Best Local Similarity	96.48;	Pred. NO. 0;		
Matches 2011; Conservative	0;	Mismatches 35;	Indels 40;	Gaps 1;

OY	3579	ctgactcgcctgcgcctcgcgttcgttcctgcgcgcgcgcgcgtatcagctcaactcaaaagcgg	3638
OY	3579	ctgactcgcctgcgcctcgcgttcgttcctgcgcgcgcgcgcgtatcagctcaactcaaaagcgg	3638
Db	2056	CTCAGCGCGTGGCGCTGGCTGTTGGCGCTGGCGGCGAGCGCGATACCTCACTCAAAAGCGCG	1997
OY	3639	taatacggcttataccacagaatcaggaggatatacgcaggaagaacatgtatgacaaaagggc	3698
Db	1996	TATATACGGTTATCCACAGAAATACGGGGATACGCCAGAAAAGAACATGTATGCAAAAAGGCC	1937
OY	3699	agcaaaagccaggaaccgtaaaaaagccgcgtctgcgtgcgttttccataagcttcgcgc	3758
Db	1936	AGCAAAAAGGCCAGGAACCGTAAAAAAGCGCGCGTGGCTGGCGTTTTCATATAGCTTCGGCC	1877
OY	3759	ccccctgaagagatacacaanaaaatgcagcttaagtctaaagcttgaggagaaacccgcagagac	3818
Db	1876	CCCCCTGACGAGATACACAAAATTCAGACGCTCAAGTCAAGAGGTGGGCAAAACCCGACAGGAC	1817
OY	3819	tataaagatacaggcgcgtttcccccctggaagctccctctgtgcctctccctgttccgaacc	3878
Db	1816	TATAAAGATACCAAGGCGTTTCCCGCTGGAAAGCTCCCTCGTGGCGCTCTCCTGTTCGACCC	1757
OY	3879	tgcgccttaacggatacctctgcgcctcttcccttcgggaacgctlgcgtcttccata	3938
Db	1756	TGCCCTTACCCGGAATACGCTCCGCCCTTCTCCCTTCCGGGAAGCGTGGCGCTTTCCAAT	1697
OY	3939	gctcagcctgtatgatatcctcagtttcggttgagtctgtctgcacaaagctggcctgtgtac	3998
Db	1696	GCTCACGCTGTAGGTATCTCACTTCCGCTTAAGTGCCTTGCTCTCAAAAGCTGGCGCTGTGTC	1637
OY	3999	acgaaaccccccgcttaagcccgacgcgctgcgccttatacgcgtiaactatcgttctgaatcca	4058
Db	1636	ACGAACCCCCCGTTCAGCCCGACCGCTGGCGCTTATTCGGTAATCATATCGTGTGAATCCA	1577
OY	4059	accgcgttaagaacagactatctgcacactgycacgcgcgcgcgcacactggtlaacagatlaagaag	4118
Db	1576	ACCCCGTAAGACACGACTTATCGCACCTGGACACCACTGGTAACACAGATTAGACAGAG	1517
OY	4119	cgagttatgagcggtgtctacagatctcttgaagtgtygccttaactacgctacacta	4178
Db	1516	CGAGGTATATAGCGCGTGTCTACAGATTTTGAAGTGGTGGCTTAACATACGCTACACTA	1457
OY	4179	gaagacagattctgttaltcgcgtctgcgtctgaagccagttaccttcggaaaaaagcttg	4238
Db	1456	GAGGACACATATTTGGTATCTCGCGCTCTCTGAAGCACTTACCTTCGGAAGAAAGAGTTG	1397

QY	4239	gtagctcttgatccggcaacaaccaaccgctggtgtagcggtgttcttctgttgcaagc	4288
Db	1366	GTAGCTCTTGATCCGGCAACAACAACCGCTGTGTAGCGGTGTTTTTGTGTTCGAAGC	1357
QY	4299	agcagattacgcgcgaagaaaaaagggtctccaagaagaaccttgatctcttcttaagggct	4358
Db	1336	AGCAGATTACGCGGAGAAAAAGATCTCCAAAGATCTTTTGATCTTTCACGGGGT	1277
QY	4359	ctgagctcgaatgggaagaaacatcaagtaagggaatttggcatalgagatatacaaaa	4418
Db	1276	CTGAGCGCTGAGTGAACCTAGATC-----	1252
QY	4419	ggatcttcaacttagatccctcttggggggggggcgctgaagtcctgcctgtgaagaag	4478
Db	1253	--GATTAAATTAAACCTTCGAATTCCTGAGCTGACTGTAGGTCTGCCTCGTGAAGAAGG	1197
QY	4479	tgttgtgactacataccaagccctgaatcgccccaatacccaacgaagaagtgaaggagcc	4538
Db	1136	TGTTGTGACTACATACCAAGCGCTGAATCGCCCATATCCACGACGAAAGTGAAGGAGCC	1137
QY	4539	acggtctgaatgagcctctgtctgtagtbtgaacagatcttgatcttgaactcttctgc	4598
Db	1136	ACGGTGTGATGAGACCTTGTGTGTAGGTGAGCAGCTGTGTGATTTTGACTTTGTCTTGC	1077
QY	4599	cacggaacgctctgcgtctgtcgggaagaatgcgtgatctgaatccctcaactcaagcaaat	4658
Db	1076	CACGGAAACGGCTCTCGTTGTTCGGGAAGATCCGTGATCTGATCTTCAACTCAGCAAAAGT	1017
QY	4659	tcgatttatcaaaaagccgcgcgtccgcgtcaagtgaagtgatgctctgcgaagtac	4718
Db	1016	TGCGTTTATTTCAAAAGCCGCGCTCCGCTCAAGTACGGTAAATCTGCGCAGGTATAC	957
QY	4719	aacctaatcaacaaattcttgattgaagaaactaactgaacataaigtgaactcaatla	4778
Db	956	AACCAATTAAACCAATTTCTGATTAGAAAAAATCTACGACATCAAAATGAACCTGCATTTA	897
QY	4779	tccatalcagaatatatacaatacatattttgaaaaagccgttctcgtlaatgaaggaa	4838
Db	896	TTCAATATCAGGATTTATCAATTAACCATTTTGTAAAAAGCCGTTTCTGTAAATGAAGCAA	837
QY	4839	aactcacggagcgagcttccatagatgagcaagatcctgttatcggctctgagatccgact	4898
Db	836	AACTCACGGAGCGAGTTCATATAGATGCAAAATCTCGTATCTGCTCGCATTTCCGACT	777
QY	4899	cgccaacaatacaaaccttatatttccccctgcaaaaaataagttatcaagttag	4958
Db	776	CGTCCAAATATAATCAACCTATTAAATTTCCCTCGTCAAAAATAAGTTATCAAGTAG	717
QY	4959	aaatcaccaatgagtgcagactgaaatccggtgtagaaltgycgaagaagctatgatctcttc	5018
Db	716	AAATCACCATGAGTAGAGACTGAATCCGGTAGAATAGCAAAAGATTATGACTTTCTTTC	657
QY	5019	cagacttgttcaaaagcgagccatcagctctgtctatcaaaatcactgcgatacaacaa	5078
Db	656	CAGACTTGTTTAAACAGGCCAACCCATTAAACGCTGCTATCAAAATCACTGCCATCAACAA	597
QY	5079	ccgttatcatctctgttatcttgcgcctgagcgagacggaataacgcatcgcgtgtctaaaga	5138
Db	596	CCGTTATTTCATTTCTGTGATTGCGCTGAGCAGACGAAATACCGGATCGCTGTTAAAGGA	537
QY	5139	caatcacaaacaggaatcgaaatgcaacccggcggaaggaacacttgcagcgatcaacaa	5198
Db	536	CAATTACAAACAGGAATCGAATGCAACCGCGCAGGAACACTGCCGATCAACAMTA	477
QY	5199	ttttcaacttgaaatcaagatatctcttaataactcgaatgcgttttcccgagatgcga	5258
Db	476	TTTTTACCTGGAATCAGGATATTCTTTCTAAATACCTGGAATGCTGTTTTCCCGGGGATCCCA	417
QY	5259	gttgtgtatcaacatgatacaacgaaggtacggaataaaatgctatagtgtctggaagagc	5318
Db	416	GTGTGTGATTAACCATGCTATCTACGGAGTACGGATAAATGCTGTGATGTCTCGAAGAGGC	357
QY	5319	ataaatcttcgcacgcagatttagctctbacacatctctgtataacatcttgacaacgcga	5378

QY	4719	aaccatataaccaattctctgatactagaagaatactcagacatcaaatgaataactgcattta	4778
DB	1016	TCGATTATTTCACAAAGAGCCGCGCCGTCACAAAGTAAAGGTAAATGCTCTGGCAGAGTTAC	957
QY	4779	ttcatatacgaatatatacaatacatatttttgaaaaagccgtttctctgtaatgaagagaa	4838
DB	896	TTTCATATCAGGATTTATCAATACCATATTATTGAAAAAGCCGTTTCTGTAAATCAAGCAAA	837
QY	4839	aactcacccgaagcagcttcacatagatbgaagaatccctgttatcgtctcgtcgtatccgact	4898
DB	836	AACTCACCGAGGCGATTCCATAGAGATGGCAAGATCCTGGATATCGGATCGCATTCGCACT	777
QY	4899	gltccaacatacaatacaactatataatttccctctgtcaaaaaataagttatcaagttag	4958
DB	776	CGTCCAAATCAATAAACCAACCTATTAAATTTCCCTCGTCAAAAATAAGTTATCAAGTAG	717
QY	4959	aatcacacatagatgagcagactgaaatcccggtgagaatbgaagaagcttatgatactcttc	5018
DB	716	AAATCACCATGATGAGACACTGATCCGGAGAAATGGCAAAAGATTATGCAATTTCTTTC	657
QY	5019	cagacttgttcaaaagccgaagccatcagctctgtatcaaaaatcctcgtcatcaacaa	5078
DB	656	CAGACTTTTTCACAGGCGACCCATTACGCTCGTCAACAAATCACTGCATCAACACAA	597
QY	5079	ccgtatcatcctctgtatctgcgcctgtagcgaacgaataacgcgtatcgtataaaga	5138
DB	596	CCGTTATTTCATTCGGATGGGCGCTGAGCAGCAAAATACGGAGATCGCTGTAAAGA	537
QY	5139	caattacaagaagaatcgaatgaaacccgtagaagaataacgtccagagcatcaacata	5198
DB	536	CAATTACAAACGGAATTCGAATGCAACCGCGCAGAACCTCCACGCCATCAACATA	477
QY	5199	ttttcaacctgaatcagaatactctctcaactacctggaatgctgttttcccgaggatcga	5258
DB	476	TTTTCACTCGAATCAGATATTCTTTTAATACCTGGAAATGCTGTTTTCCGGGATGCCA	417
QY	5259	gtgtgtgaaataacatgcatactcaagagatcggataaataatgcttgaatgtgcgaagagc	5318
DB	416	GTGGGAGATTAACATGCATCTCACTGAGGTACGGATTAATAATGCTGTAGTGTCCGAAGAGCG	357
QY	5319	ataaatctcgtcagccagtttagctctacacatctcatctgataacatcatctggacaagcta	5378



D <sub>b</sub>	356	ATAAATTCCGTCGCCAGTTTATGTCTGCACCATCTATCTGTAAACTCAATTGGCAACGCTA	297
Q <sub>y</sub>	5379	ccttgcacatggttcagaacaactctcgtcgatcgagtgctccatacatcgatagt	5438
D <sub>b</sub>	296	CCTTGCGCATGTTTCAGAAMAACAACCTCGCGCAGTAGGGCTTCCCCTAACAGGAATAGATT	237
Q <sub>y</sub>	5439	gtcgaccctgatggcccgaatatattcgtgaagccattatcacatataaatcaacatcc	5498
D <sub>b</sub>	236	GTCGCACCTGATTGCCCGAATTAATGCGCGGCCAATTTATACCCTATTAATCACACATCC	177
Q <sub>y</sub>	5499	atgttgttaattaaatcgcgcgcctcgcgaagaagttcccgltgaatlatgctcataca	5555
D <sub>b</sub>	176	ATGTTGGAAITTTAATCGCGGCGCTCGAGCAAGACGTTTCCCGTTGAATAATGGCTCATACA	117
Q <sub>y</sub>	5559	ccccctgataactggtttaatgtaaagcagaagaagtttatgttctaatagatataattta	5616
D <sub>b</sub>	116	CCCCGTGATTAACGTATTATGTAAGCGAAGAAGTTTATGTTTCATATGATATATTTTTTA	57
Q <sub>y</sub>	5619	tctgttgcaatgtaaacatcaagatattcttgagaacaacgtgacctt	5664
D <sub>b</sub>	56	TCTTGTGCAATCTAATCAATCAAGATTTTGAGACACAACGTCGCTTT	11







Dp	1725	CGACCCTGGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTT	1663
Oy	3933	ctcatagctcagcgtctgaagtatcttcagtttcgtgtgaagtgcgtccaaagcttggc	3992
Dp	1665	CTCAATGCTTCACGCTGTAGGTATCTCAGTTGGGTGAAGTGTTCGCTCCAAAGCTGGGCT	1606
Oy	3993	gtctgcacgaacccccgcgtctcagcccgacgcgtctgccttaaccgtaacatcgtctg	4052
Dp	1605	GTGTGACGACGAACCCCCGTTCAACCCGACCGCGTGGCGCTTATCCGGTAATATGCTGTG	1546
Oy	4053	agttcaaccccggttaagaacacgactatcgccaccttgcagcagcagccactgttaacaggtat	4112
Dp	1545	AGTCCAAACCCGGTAACACACGACTTATCCCACTGGCAACACCCACTGTTAAACGGATTA	1486
Oy	4113	gcagagcgaaagtctgttagagcggtgtgcacagaaagttcttcaagtgtgtgcctaaacgct	4172
Dp	1485	GCAAGCGGAGGTATGTAGCGCGGTCTACAGAGTTCTTGAAGTGTGTGGCTTAACTAACGGCT	1426
Oy	4173	acactagaagaaagatatttgttatctgcgtctgtcgtgaagccagttaccttcgnaaaa	4232
Dp	1425	ACACTAGAAAGACAGTATTTGGTATCTGGCGCTGTGCTGAAGCAATTAACCTTCGGAAAA	1366
Oy	4233	gacttggtagctcttgaatcccggaacaaacacacccgctgtgtgacgtgtgttttggtt	4292
Dp	1365	GAGTGTGTACTCTTGATCCGGCAAAACCAACCGCTGTAGCGGTGGTGTGTTTTTGTT	1306
Oy	4293	gcaagcagcagatatacgcgcagaaaaaagatctcaagaagaatccttgaatcttctta	4352
Dp	1305	GCAAGCAGCAGATATTACCGCGCAAAAAAAGATCTCAAGATGTCCTTGTGATCTTTCTA	1246
Oy	4353	cggtgtctgacgtcagttgaacgnaaaactcaagtttaaggaatttggltcatgagttat	4412
Dp	1245	CGGGGTGTGACGCTCACTGGAACCTA-----	1219
Oy	4413	caaaaagatctcaactagatccttttggggggggggcgctgagttcgtccctcgtga	4472
Dp	1220	-----GTAGAGTGTGCTCCTCGGA	1209
Oy	4473	agaagttgttgcgtactacataccaagccttgaatcgcccatcatccaagcagaagaatgag	4533
Dp	1202	AGAAAGTGTTCCTGACTCATATACAGAGCGCTGAATCGCCCATCTACGACGACAAGATGAG	1143
Oy	4533	ggaagccaggttgaagaagcttgtgtgtagtggagccagtttggatatttgaactttg	4592
Dp	1142	GGAGCCACGGTGTATGAAAGCTTTGTTAGTGGGACCACTTGATTTTAACCTTTTG	108
Oy	4593	ctttgcacagaaagcgtctcgttctgcggaagaatgcgtatctgaacctcaaacacagc	4653
Dp	1082	CTTTGCCACGGAAAGGTCGTGCTTGTGGGGAAGATGGATCTGATCTTCAATCAACAC	1022
Oy	4653	aaaagttcgatttatcaacaagaacgcgctcccgctcaagtgaagttaatgtcttcgcaag	4713
Dp	1022	AAAATGTGATTATTATCAACAAACCGCGCTCCCGTCAATCGCTAATATGCTCTGCGCAG	963
Oy	4713	tgttcaacccaattaaaccaatttcgtatagaaaaaacatcatcgcagatcaaalgaaatcgc	4772
Dp	962	TGTTACACCAATTATACCAATTTCGATTAGAAAAACATCGAGTCATCAATGAATGAACCTGC	903
Oy	4773	aattattcataccaagattatacaataccatatatttgaaaaaagccggttcgttaatga	4833
Dp	902	AATTATTTCATATACGAGATTATCAATATCCATATTTTGGAAAAAGCCGTTCTGTATGAAA	843
Oy	4833	ggaagaanaactaacaggaagcattcatatagatgtgcaagatccctggtatcgtctgcat	4892
Dp	842	GGAGAAAATCTACCGAGGCACTTCCATATAGATGGCAAGATCCTGTGTATCGGTGTGGATT	783
Oy	4893	ccgactcgtccaacatacaaacactatnaatttccctcgtccaanaaataaggtatca	4953
Dp	782	CCGACTCGTCCAACTCAATCAACACTATTAATTTCCCTCGTCAAAAATAAGATTATCA	723
Oy	4953	agtgtgaatatcaaccatgtgtgtgaagctgaatccggttgaagatgtgcnaaagcttatcat	5013
Dp	722	AATTGGAATATCCATGTAGGTATACGACTGAATTCGGTGTAGATGGCAAAAGATTATCATTT	663

QY	5013	tccttcacagcttggtccaagaagccagccattacgctcgtctcaaaatcactcgtcatca	5072
Db	662	tccttcacagcttggtccaagaagccagccattacgctcgtctcaaaatcactcgtcatca	603
QY	5073	accaaacggtatctctcgtgattcgcgcctcagcagaaacgaatacagcatcgctgta	5132
Db	602	accaaacggtatctctcgtgattcgcgcctcagcagaaacgaatacagcatcgctgta	543
QY	5133	aaaggaacaattacaacaggaatcgaatgcgaacccgcgcgaagaaacacgccaagcgcatca	5192
Db	542	aaaggaacaattacaacaggaatcgaatgcgaacccgcgcgaagaaacacgccaagcgcatca	483
QY	5193	acaatattccactgaatcaagatactcttcaatccactcgaatcgcgtttcccgagg	5252
Db	482	acaatatttttccactgaatcaagatattcttcaatccactcgaatcgcgtttttcccgagg	423
QY	5253	atcgcagtggtgagtaaccatgcatacgaagaagtaacgaataatgccttgatgctcga	5312
Db	422	atccgactggtagatgaaccatgcatacgaagaagtaacgaataatgccttgatgctcga	363
QY	5313	agaaagcaataattccctgcagccagtttagctcgcacatctcgtcgttaacatcatctgcga	5372
Db	362	agaaagcaataattccctgcagccagtttagctcgcacatctcgttaacatcatctggca	303
QY	5373	acgctaccttgcgcatcttccaagaacacatcgcgcacatcgcgccttcccatcaatcga	5433
Db	302	acgctaccttggcgatcttccatgcttttcagaaacaaactcgtggcgatggggctttcccatcaacgca	243
QY	5433	tagatgtcgcgaacctgattgcgcgaacattatcgcgagcccaattataccatataatca	5493
Db	242	tagatgttcgcgacctgattgcgcgaacattatcgcgagcccaattataccatataatca	183
QY	5493	gcacccactgttggaattaatcgcgcctcgcgaagaacgcttcccgctggaatagctc	5552
Db	182	gcattccattgttggaattaatcgcgcctcgcgaagaacgcttcccgctggaatagcgctc	123
QY	5553	ataaacacccttgattacttgattatgtaagcagacaagtttatgtgtcattgatgatata	5612
Db	122	ataaacacccttgattacttgattatgtaagcagacaagtttatgtgtcattgatgatata	63
QY	5613	ttttattctgtgcgaatgtaacatcagaagtttttgagaacacagcgagctt 5664	
Db	62	tttttatctgttgcaatgtaacatcagaagtttttgagaacacagcgagcttt 11	
RESULT 12			
PCT-US95-12642-14/c			
Sequence 14, Application PC/TUS9512642			
GENERAL INFORMATION:			
APPLICANT: Pathogenesis Corporation			
APPLICANT: 201 Elliott Avenue West, Suite 150			
APPLICANT: Seattle, Washington 98119			
TITLE OF INVENTION: Mycobacterial Reporter Strains and Uses			
NUMBER OF SEQUENCES: 18			
CORRESPONDENCE ADDRESS:			
ADDRESS: Townsend and Townsend Khourie and Crew			
STREET: Steuart Street Tower, One Market Plaza			
CITY: San Francisco			
STATE: California			
COUNTRY: US			
ZIP: 94105-1493			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: PCT/US95/12642			
FILING DATE: 02-OCT-1995			
CLASSIFICATION:			
PRIOR APPLICATION DATA:			

RESULT 12  
 PCT-US95-12642-14/c  
 : Sequence 14, Application PC/US9512642  
 :  
 GENERAL INFORMATION:  
 :  
 APPLICANT: Pathogenesis Corporation  
 APPLICANT: 201 Elliott Avenue West, Suite 150  
 APPLICANT: Seattle, Washington 98119  
 TITLE OF INVENTION: Mycobacterial Reporter Strains and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSSEE: Townsend and Townsend Khourie and Crew  
 STREET: Steuart Street Tower, One Market Plaza  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: US  
 ZIP: 94105-1493  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/12642  
 FILING DATE: 02-OCT-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:























Db 233 GCACCTGATTGCCGACATTAACGCGAGCCCATTTATACCCATTAATAACAGCATCCATG 174  
QY 5502 ttggaatttaacgcgggctcgaagaagcgttcccgltgaalatygcloataaacacc 5561  
Db 173 TTGGAATTTAATCGCGGCTCGAGCAGACAGCTTCCCTTGATATATGCGCTCATACACCC 114  
QY 5562 ctgtatactgttlatgtaagcagacagtttattgtlcatgaltatattttatct 5621  
Db 113 CTGTATTAAGTGTATGTAAGCAGACAGTTTATTTGTCATGATGATATATTTTATCT 54  
QY 5622 tgtgcaatgtaacatcagagatttggagacacacacgtggtt 5664  
Db 53 TGTCAATGTACATCAGAGATTTTGAGACACACGTCGCTTT 11

Search completed: August 20, 2002, 20:25:59  
Job time: 9208 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 17:51:26 ; Search time 4096.96 Seconds  
(Without alignments)  
19301.790 Million cell updates/sec

Title: US-09-826-206-3  
Perfect score: 5859  
Sequence: 1 tcgcgcgttcggtgatgc.....taccagagcccttcgctc 5859

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estl:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vit:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	743.8	12.7	800	AJ281449	AJ281449 4A3A-P4D5
2	736.4	12.6	1004	AJ281449	AJ281449 4A3A-P4D5
3	722.4	12.3	724	AL645114	AL645114 4A3A-P4D5
4	719.2	12.3	988	CNS06X6V	AL419357 77 end of
5	712.4	12.2	954	AL044364	DKFZP434C
6	711.8	12.1	1067	AJ281449	AJ281449 4A3A-P4D5
7	708.4	12.1	1070	AJ281552	4A3A-P4D5
8	698	11.9	718	AL631067	AL631067 4A3A-P4D5
9	680.6	11.6	700	AJ281616	AJ281616 4A3A-P4D5
10	680.4	11.6	1089	AJ281124	AJ281124 4A3A-P4D5
11	672.4	11.5	675	AL636713	AL636713 4A3A-P4D5
12	665.8	11.4	672	AL640799	AL640799 4A3A-P4D5
13	664.8	11.3	671	AL660035	AL660035 4A3A-P4D5
14	662.4	11.3	675	AL645322	AL645322 4A3A-P4D5
15	662.2	11.3	670	AL642581	AL642581 4A3A-P4D5
16	651.2	11.1	668	AL637190	AL637190 4A3A-P4D5
17	645.4	11.0	658	AL654693	AL654693 4A3A-P4D5

18	645.4	11.0	659	AL647428	AL647428 4A3A-P4D5
19	644.4	11.0	658	AL656146	AL656146 4A3A-P4D5
20	644.4	11.0	689	AL646532	AL646532 4A3A-P4D5
21	644.4	11.0	707	AL656688	AL656688 4A3A-P4D5
22	643.6	11.0	808	AJ176264	AJ176264 4A3A-P4D5
23	643.4	11.0	657	AL641508	AL641508 4A3A-P4D5
24	643.4	11.0	670	AL635952	AL635952 4A3A-P4D5
25	642.8	11.0	705	AL643164	AL643164 4A3A-P4D5
26	642.8	11.0	715	AL661706	AL661706 4A3A-P4D5
27	641.4	10.9	669	AL660789	AL660789 4A3A-P4D5
28	640.8	10.9	705	AL635845	AL635845 4A3A-P4D5
29	639.8	10.9	651	AL660958	AL660958 4A3A-P4D5
30	636.2	10.9	671	AL649526	AL649526 4A3A-P4D5
31	635.6	10.8	718	AG105205	AG105205 Pan trogl
32	635.2	10.8	719	AG089715	AG089715 Pan trogl
33	631.8	10.8	646	AL633911	AL633911 AL633911
34	624.6	10.7	734	AL039459	DKFZP4340
35	624.4	10.7	654	AL639359	AL639359 AL639359
36	622.4	10.6	647	BH241665	BH241665 ATG0827R
37	618.8	10.6	703	AJ281437	AJ281437 4A3A-P4C3
38	618.4	10.6	718	AG010489	AG010489 Homo sapi
39	614.4	10.5	638	AL642392	AL642392 AL642392
40	613.2	10.5	695	AG100045	AG100045 Pan trogl
41	612	10.4	632	AL638320	AL638320 AL638320
42	609.4	10.4	664	AL650636	AL650636 AL650636
43	609.4	10.4	688	AG095573	AG095573 Pan trogl
44	609	10.4	706	AG085503	AG085503 Pan trogl
45	608.8	10.4	626	AL634940	AL634940 AL634940

## ALIGNMENTS

RESULT	1	800 bp	LINEAR	EST 30-JUN-2000
AJ281449	4A3A-P4D5-F	Anopheles gambiae	immune competent	4A3A Anopheles
LOCUS	gambiae cDNA clone 4A3A-P4D5	800 bp	LINEAR	EST 30-JUN-2000
DEFINITION	gambiae cDNA clone 4A3A-P4D5	800 bp	LINEAR	EST 30-JUN-2000
ACCESSION	AJ281449	GI:6929329		
VERSION	AJ281449.1	GI:6929329		
KEYWORDS	EST			
SOURCE	African malaria mosquito.			
ORGANISM	Anopheles gambiae			
REFERENCE	1 (bases 1 to 800)			
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansong, W., Soares, M. B. and Kafatos, F. C.			
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12)	6619-6624	(2000)
MDLINE	Contact: Dimopoulos G			
COMMENT	Fotis C. Kafatos laboratory			
FEATURES	European Molecular Biology Laboratory			
source	Meyernotstrasse 1, 69117 Heidelberg, Germany.			
	Location/Qualifiers			
	1..800			
	/organism="Anopheles gambiae"			
	/strain="4A r/r"			
	/db_xref="taxon:7165"			
	/clone="4A3A-P4D5"			
	/clone_id="Anopheles gambiae immune competent 4A3A"			
	/cell_line="Immune competent 4A3A"			
	/lab_host="E. coli DH10B"			
	/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line			







[illegible]

Db	61	AAAGCGGTAATACGCTTATCCACGAAATCAAGGAGGATTAACCGAGGAAGAAGCAATGATGAGC	120
QY	3691	aaagagccagcaaaaagccaggaacccgtaaaaaagccgctgtgcgtgtttccatag	3750
Db	121	AAAAGGCCAGCAAAAAGGCCAGAACCCGTAAAAAAGGCCCGCTGGCGCTTTTTCATAG	180
QY	3751	gctccgcccccttgacgagatcacaanaatcgacgctcaagtcaagggtgacgaacc	3810
Db	181	GCTCCGCCCCCTGACGAGCATCAACAAAATCGACGCTCAAGTCAGAGGTGGCAAAACC	240
QY	3811	gacagactataaagataccaggcgtttccccctggaagctcccctcgtgcctcctgt	3870
Db	241	GACAGAGCTATTAAGATACAGAGCGCTTTCCCTTGGAAGCTCCCTCGGCCCTCTCTGT	300
QY	3871	tcagacctgcgcgtctaccagatctctgcgcctttcccttccttcggaaagcttggcct	3930
Db	301	TCCGACCTGCGCGCTTACCGGATFACCTGTCCGCTTTCTCTCCCTTGGGAAGGTGCGCT	360
QY	3931	ttctcatagctcaacgctgtagtactcagttcagttcagttgtagtgcgttcgctcaagctg	3990
Db	361	TTCCTATFAGCTCAGCCTGAGGTATCTCAGTTCCGGTGAAGTCTGCTCCCAAGCTGG	420
QY	3991	ctgtgtgcagcaaccccccgcttcagcccgacccgctgcgcttatccggtaatatgctc	4050
Db	421	CTGTGTGACGAAACCCCGCTTACCCGACCGCGCTGCCCTTATTCGGTACTATACGCT	480
QY	4051	tgaagtcacacccggtgaagaacagactttatgcacattgcaagcagcactgtgtaagat	4110
Db	481	TGAGTTCACACCCGGTAAAGACACGACTTATCCGCACTGGCAGCAGCCACTGGTAACAGAT	540
QY	4111	tacgagagcagagtagtgaagcgtgtcctacagagttcttgaagtgggtgccaactacag	4170
Db	541	TAGCAGACGAGAGTGTAGCGCGGTCAACAGATTCCTTGAAGTGGTGCTTAACACG	600
QY	4171	ctaacctgaagaagacagatatttgaattgtgcgtctgtcttgtaagccagttactctgaa	4230
Db	601	CTACCTGTAAGAAACAGATTGTGGATCTGGCGCTCTGCTGAAGCCAGTTACTCTTGGA	660
QY	4231	aagaattgtagctcttcatccggaacaaacacccgctgtagcggtagtlttttgt	4290
Db	661	AAGAGTGGTACTCTTGATCGCGCAACAAACACCGCTGTGACGGTGGTTTTTTTGT	720
QY	4291	ttgc 4294	
Db	721	TTGC 724	
RESULT	4		
CNS06X6V			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			











REFERENCES  
AUTHORS

TITLE

JOURNAL  
MEDLINE  
COMMENTFEATURES  
SOUND

```

SOURCE
/organism="Anopheles gambiae"
/strain="4A r/Y"
/db_xref="taxon:7165"
/clone="4A3A-p6F11"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="vector: pRTT3D-pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
c1190-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two Approaches To
Facilitate Gene discovery, Genome Research 6, 791-806."

```

Query Match	12.1%;	Score 708.4;	DB 9;	Length 1070;
Best local Similarity	99.9%;	Pred. No. 4e-162;		
Matches 709; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	4211	aagcagcttaacctcggaagaaagatgtgtgaacttgaatccggaaaacaacccgct	4270
Db	481	AAGCCAGTTCACCTTCGGAAAAAGATTGGTACCTCTTGATCCGGCAACCAACCCCT	540
QY	4271	gttagcgtagtattttttgtttgttgaacgacgagatctacgcgcagaaaaaagatctaa	4330
Db	541	GGTAGCGGTGGTTTTTTTGTTCAGACAGACATATACCGGAGAAAAAAGATCTCA	600
QY	4331	gaagatcccttgatcttttcttaacggggtcgaagcctcagtgaaacgaaaaccacgttaa	4390
Db	601	GAACATCCCTTGATCTTTTCTACGGGGTCTGAAGCCTCAATGAAACCAAAACCTCAGTTAA	660
QY	4391	gggatttgcacgagataltcaaaaagatcttcaactagaatcccttt	4440
Db	661	GGGATTTGGTCATGAGATTATCAAAAGATCTTCACTAGATCCTTTT	710

	RESULT	8						
	LOCUS	AL631067						
	DEFINITION	AL631067 XGC-gastrula Silurana tropicalis cDNA clone Tga015cell 5'	718 bp	mRNA	linear	EST 02-NOV-2001		
		nrna sequence.						
	ACCESSION	AL631067						
	VERSION	AL631067.1	GI:16600550					
	KEYWORDS	EST.	western clawed frog.					
	SOURCE	Silurana tropicalis	Eumecyria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	ORGANISM	Ampibiia; Batrachia; Anura; Mesobatrachia; Pipidoes; Piploae; Xenopodinae; Silurana.	1 (bases 1 to 718)					
	REFERENCE	Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.						
	AUTHORS	Sanger Xenopus tropicalis EST project 2001 (10-2001)						
	TITLE	Unpublished (2001)						
	JOURNAL	Contact: Huckle E						
	COMMENT							

```

FEATURES
  source
    Location/Qualifiers
      1..718
        /organism="Silurana tropicalis"
        /db_xref="taxon:8364"
        /clone="Tgas015e11"
        /clone_lib="XGC-gastrula"
        /dev_stage="gastrula (stages 10.5-13 mixed)"
        /lab_host="Escherichia coli XL1-blue"
        /note="Vector: PCS107; Site.1: EcoRI; Site.2: NotI; cDNA
        was oligo dT primed from 5'ug of poly A+ RNA from stages
        10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
        into PCS107 with EcoRI at the 5' end and NotI at the 3'
        end."
BASE COUNT      172 a      197 c      187 g      162 t
ORIGIN
Query Match      11.9%; Score 698; DB 9; Length 718;
Best Local Similarity      Pred. No. 1,2e-159;
Matches 713; Conservative      0; Mismatches 5; Indels 2; Gaps 1.

```











Qy 3807 acccgacagactataaagataccaggcggttcccccgtgaagctccctcg tgcgtctc 3866  
|||||  
Db 182 ACCGACAGGACTATAAGATACAGAGGGTTTCCCGCGAAGCTCCCTCGGCTCTC 241  
Qy 3867 ctgttcgacccctgcgcttaacggatacctgtccgccttcccttcgggaacgctg 3926  
|||||  
Db 242 CTGTTCCGACCTCCCTTACCGGATACCTGTCCGCTTCTCCCTTCGCGAAGCGTGG 301  
Qy 3927 cgcttccatagctacgcgtgtatctcagttcggtgtaggtgttcgtcccaagc 3986  
|||||  
Db 302 CGCTTCTCAATAGCTCAAGCTGATGATCTCAGTTGGGTAGGTGCTTGCCTCCAGC 361  
Qy 3987 tggagctgtcaagaaaccccgcttaagccgacgctgtgccttaacgtaacttc 4046  
|||||  
Db 362 TGGGCTGTGTGTCAGAACCCCGCTTACGCCGACCGTGGGCTTATCCGTAATCTC 421  
Qy 4047 gcttgaagtcacacccggttaagacacgactatcgccacttgcagcagcactgttaaca 4106  
|||||  
Db 422 GTCTTGAGTCCAAACCCGGTAAGACAGCACTTATCGCCACTGGACAGCCACTGGTACA 481  
Qy 4107 ggatlaacagaggaagtatgtagcggtgtctacagagttcttgaagtgtgtgcttaact 4166  
|||||  
Db 482 GGATTACAGAGCGAGGATGTAGCGCGGTCTACAGAGTTCTTGAAGTGAGGCTTAATCT 541  
Qy 4167 acgagctacacagagaagacagatattgtatctgcgctctgtgaagccagtiacctcg 4226  
|||||  
Db 542 ACGGCTACACTAGAAAGAACAGTATTGGTATCTGCGCTCTCTGTAAGCAATTAATCTTCG 601  
Qy 4227 gaaaaagagtgtagctctgtatctgcgcaacaacacccgctgtgtagcggtgttlt 4286  
|||||  
Db 602 GAAAAAGAGTGTAGCTCTTGTATCGGCAAAACACCGGTGTAGCGGTGTGTTT 661  
Qy 4287 ttgtttgcaagcag 4300  
|||||  
Db 662 TTGTTTGCAGACAG 675

RESULT 12  
AL640799 672 bp mRNA linear EST 12-DEC-2001  
LOCUS AL640799 XGC-neurula Silurana tropicalis cDNA clone TNeu005c22 5',  
DEFINITION mRNA sequence.  
ACCESSION AL640799  
VERSION AL640799.1 GI:16792930  
KEYWORDS EST.  
SOURCE western clawed frog.  
ORGANISM Silurana tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Silurana.  
REFERENCE 1 (bases 1 to 672)  
AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.  
TITLE Sanger Xenopus tropicalis EST project 2001 (10\_2001)  
JOURNAL Unpublished (2001)  
COMMENT Contact: Huckle E  
Sanger Centre  
Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: tropesanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS\_SEQUENCE\_ID: TNeu005c22.sp6  
Sequencing primer: SP6  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn.  
FEATURES  
source  
1..672  
/organism="Silurana tropicalis"  
/db\_xref="taxon:8364"  
/clone="TNeu005c22"  
/clone\_lib="XGC-neurula"  
/dev\_stage="neurula"  
/lab\_host="Escherichia coli DH10B"  
/note="Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI; CDNA

BASE COUNT 152 a 187 c 173 g 160 t  
ORIGIN  
Query Match 11.4%; Score 665.8; DB 9; Length 672;  
Best Local Similarity 99.7%; Pred. No. 8,8e-152;  
Matches 667; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
was oligo dT primed from 5ng of poly A+ RNA from neurula.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with  
EcoRI at the 5' end and NotI at the 3' end.  
Qy 3720 aaaaagccggttgcgttgcgttcttccataggtctcccccgtgacagcatcaaaaa 3779  
|||||  
Db 4 AAAAAAGCGCGTGTGTGCGGCTTTTTCATAGGCTCCGCCGCCCTGACGATCAAAAA 63  
Qy 3780 atcgaagctcaagatcaaggtgtgtggaacccgcagacagactataaagataccagcgcttc 3839  
|||||  
Db 64 ATCGACGCTCAAGTCAAGTGAAGTGGCGAAACCCGACAGGACTATMAAGATACAGCGGTTTC 123  
Qy 3840 ccccttggaagctccctctgtgcgtctcctgttccgaacctgcgcttaaccgataacctgt 3899  
|||||  
Db 124 CCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCAGACCTGCGGCTTACCGGATACCTGT 183  
Qy 3900 ccgaccttctcccttcggaagcggtgcgttctcatalagctcaagctgttagtata tctca 3959  
|||||  
Db 184 CCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATATAGCTCACGCTGTAGTATCTCA 243  
Qy 3960 gtctgtgttagtgtctgtctcccaagctgtgtgtgtgtgcaagaaccccgcttaaccccg 4019  
|||||  
Db 244 GTTCGCTGTAGGTCTTTCGCTTCAAGCTGGGCTGTGTGCACAAACCCCGTTACGCCCG 303  
Qy 4020 accgctgcgcttaccggtiaactatcgctctgtgtgttcaaccggttaagacacagactat 4079  
|||||  
Db 304 ACGGCTGGGCTTATCCGTAATCTATCGTCTTGAATCCAAACCCGGTAAGACAGCACTTAT 363  
Qy 4080 cggcactgagcagcagcactgttgaacagatlagcagaagcagatgtatgtgtgtctta 4139  
|||||  
Db 364 CGCCACTGGACAGCAGCACTGTGTAACAGATTAAGACAGCGAGTATGAGCGGTGCTTA 423  
Qy 4140 cagagttcttgaagt 4199  
|||||  
Db 424 CAGAGTTCTTGAAGT 483  
Qy 4200 ggcgtctgtgaagcagcttacccttcggaaaaagagttgtgtgtgtgtgtgtgtgtgtgt 4259  
|||||  
Db 484 GCGCTCTGTGTAAGCAAGTACCTTCGGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543  
Qy 4260 aaaccacgctgt 4319  
|||||  
Db 544 AAACACCGCGT 603  
Qy 4320 aagatctcaagaagaatccttgcattcttctcaagggtgtgtgtgtgtgtgtgtgtgtgtgt 4379  
|||||  
Db 604 AAGGATCTCAAGAAAGATCTTGTGATCTTGTACGGGGGTGTGACCTCAGTGTGAACGAAA 663  
Qy 4380 actcagctt 4388  
|||||  
Db 664 ACTCAGCTT 672

RESULT 13  
AL660035 671 bp mRNA linear EST 13-DEC-2001  
LOCUS AL660035 XGC-neurula Silurana tropicalis cDNA clone TNeu040c09 5',  
DEFINITION mRNA sequence.  
ACCESSION AL660035  
VERSION AL660035.1 GI:17674018  
KEYWORDS EST.  
SOURCE western clawed frog.  
ORGANISM Silurana tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Silurana.







Db 240 CGTTCGCTCCAGCTGGGCTGTGTGACGAAACCCCGTTACGCCGACCGCTCGCCTT 299  
 QY 4033 atccggtactatcgtcttgatccaaaccggttaagacacgactatcgcacactggcagc 4092  
 Db 300 ATCCGGTACTATGCTTGTAGTCCAAACCCGGTAGACAGACTTATGCGCACTGGCAGC 359  
 QY 4093 agccactgtaacagagatagcagagcaggtatgtagcggtgctacagagttcttgaa 4152  
 Db 360 AGCAGCTGTAACAGAGATTAGCAGAGGAGATATGTAGCGGCTCTACAGAGTCTTGAA 419  
 QY 4153 gtgtgtgctactaactaagcgtacactagaagagacgatttgglatctgcgtctgtgaa 4212  
 Db 420 GTGTGTGCTTACTACTACGCTACACTAGAACAGATTTGTGTGTGCTGCGCTGCTGAA 479  
 QY 4213 gcaagttactctcgaaagagatgtgtgactctgtatccggaacaaacacacgcttg 4272  
 Db 480 GCCAGTACCTTCGGAAAAAGATTGGTAGCTCTTGATCCGGCAACCAACACCGCTGG 539  
 QY 4273 taagcgtgttcttctgttgcgaagcagcagattacgcgcgaagaaagagatctcaaga 4332  
 Db 540 TAGCGGTGTTTTTTTGTGCAAGCAGCAGATTAGCGCGCAAGAAAAAGGATCTCAAGA 599  
 QY 4333 agatccttgatcttctctacggggtcctgacgtcagtggaagcaaacacacgttaag 4392  
 Db 600 AGATCCTTGTATCTTCTACGGGGCTGTGACGCTAGTGAAACGAAAACTCACGTTAAG 659  
 QY 4393 gatttggatcataga 4408  
 Db 660 GATTTTGCTCATGAGA 675

## RESULT 15

AL642581 670 bp mRNA EST 12-DEC-2001  
 LOCUS AL642581 XGC-neurula Silurana tropicalis cDNA clone TNeu007e10 5',  
 DEFINITION mRNA sequence.

ACCESSION AL642581  
 VERSION AL642581.1 GI:16794706  
 KEYWORDS EST.  
 SOURCE western clawed frog.  
 ORGANISM Silurana tropicalis

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 Xenopodinae; Silurana.

REFERENCE 1 (bases 1 to 670)  
 AUTHORS Huckle, E., Taylor, R., Ashurst, J. L., Zorn, A. M. and Rogers, J.  
 TITLE Sanger xenopus tropicalis EST project 2001 (10\_2001)  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Huckle E

Sanger Centre  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: tropesanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE\_ID: TNeu007e10.sp6  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Aaron M. Zorn.

## FEATURES

source

1..670

location/Qualifiers

/organism="Silurana tropicalis"  
 /db\_xref="taxon:8364"  
 /clone="TNeu007e10"  
 /clone\_1lb="XGC-neurula"  
 /dev\_stage="neurula"  
 /lab\_host="Escherichia coli DH10B"  
 /note="Vector: PCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA  
 was oligo dt primed from 5ug of poly A+ RNA from neurula.  
 EcoRI-NotI cut cDNA was then ligated into PCS107 with  
 EcoRI at the 5' end and NotI at the 3' end."

## BASE COUNT

153 a 187 c 172 g 158 t

Query Match 11.3%; Score 662.2; DB 9; Length 670;  
 Best Local Similarity 99.6%; Pred. No. 6,7e-151;  
 Matches 664; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3720 aaaaagccggtctgctggtcttcttcaataggtctccgccccctgaagacatcaaaaa 3779  
 Db 4 AAAAAGCCGCGCTTCTGCTGCTTTTCCATAGGCTCCGCCCCCTGACAGATCAAAAA 63  
 QY 3780 atcgagctcaagtcagagagtggtggaaccccgacagagcttaagaataacagggttct 3839  
 Db 64 ATGCAAGCTCAAGTCAGAGAGTGGGCAAAACCGACAGACTTAAAGATACAGAGCGTTTC 123  
 QY 3840 ccccttgagaagctccctctgctcctcctctctcgaacctcgccttaacagatacctgt 3899  
 Db 124 CCCCTGAGAGCTCCCTCGTGGGCTCTCTGTTCCGACCCCTCCGCTTACCGGATACCTGT 183  
 QY 3900 ccgcctctccctcctcggaagcgttgagcttctcctatagctcaagctgtatgtatctca 3959  
 Db 184 CCGCTTCTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGTAGTATCTCA 243  
 QY 3960 gttcgtgtagatgtctgtcgtccaaagctgggtctgtgtgcaagaccccgcttaagccg 4019  
 Db 244 GTTGGGTAGTGGTGTGCTTCCAAAGCTGGCTGTGTGTCAGAAACCCCGTTCAAGCCG 303  
 QY 4020 accgctgaccttaccggtactactcgtcttgagtccaaacccggttaagacagactat 4079  
 Db 304 ACCGCTGGCCTTATTCGGTAACTATGCTTGTAGTCCAAACCCGGTAAAGACAGACTTAT 363  
 QY 4080 cgcgaactgacagcagcagctgtgtacagagattagcagagcgaagttatgtaggcgtgcta 4139  
 Db 364 CGCCACTGGCAGCAGCCACTGTGTAACAGATTAGCAGACGAGAGATATGTAGCGGTGCTA 423  
 QY 4140 cagagttcttgatgtgtgtggtcctactacggtcctacacagaagagatattgttatct 4199  
 Db 424 CAGAGTCTTAAAGTGTGGCTTAACGCTTACACTAGAAAGACGATATTTGGTATCT 483  
 QY 4200 ggcctctgctgaagccaaatctacctcggaagaaagatctgttagctctgtatccgcaaac 4259  
 Db 484 GCGCTCTGCTTAAGCAGCACTTACCTTCGGAAGAAAGATTGTAAGCTTGATCCGGCAAC 543  
 QY 4260 aaaccaccgctgtgtagcgtgtgttcttctgttgcaagcagcagattacgcgcgaagaaa 4319  
 Db 544 AAACCAACCGCTGTAGCGGTGTTTGTGTTGCAACGACAGATTTACGCGCAGAAAAA 603  
 QY 4320 aagatctcaagaagatctcttgatcttcttctacggggtctgagcgtcaagtggaaacgaa 4379  
 Db 604 AAAGATCTCAAGAAAGATCTTTGATCTTTCTTACGGGGTGTGACGCTCAGTGAAACGAAA 663  
 QY 4380 actcaag 4386  
 Db 664 ACTCAGC 670

Search completed: August 20, 2002, 20:18:55  
 Job time: 8849 sec



**THIS PAGE BLANK (USPTO)**

Wed Aug 21 11:36:58 2002

us-09-826-206-3.rst



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 06:42:32 ; Search time 9889.5 Seconds  
(without alignments)  
18005.349 Million cell updates/sec

Title: US-09-826-206-2  
Perfect score: 8509  
Sequence: 1 tsgtgaacgacatgctgtga.....tatgttgcacgcgcacga 8509

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_pro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlgc\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	8509	100.0	8509	6	AX281574	AX281574 Sequence
2	8509	100.0	8509	6	AX281868	AX281868 Sequence
3	2517	29.6	2517	6	AX281576	AX281576 Sequence
4	2007	23.6	2112	6	AX281573	AX281573 Sequence
5	2007	23.6	5859	6	AX281575	AX281575 Sequence
6	165.6	1.9	23457	1	AE008913	AE008913 Salmonella
7	164.8	1.9	88037	1	AL627284	AL627284 Salmonella
8	122.2	1.4	10661	1	AE005851	AE005851 Caulobacter
9	92.4	1.1	12903	1	AE007289	AE007289 Stenotrophomonas
10	84.6	1.0	101711	2	AC098576	AC098576 Drosophila
11	68.4	0.8	125020	9	AF429315	AF429315 Homo sapi
12	65	0.8	9973	1	AE000070	AE000070 Rhizobium
13	65	0.8	320040	6	A79350	A79350 Sequence 1
14	65	0.8	320040	6	AF429315	AF429315 Homo sapi
15	62.6	0.7	125020	9	AF429315	AF429315 Homo sapi
16	57.4	0.7	206479	1	AB016260	AB016260 Agrobacterium
17	55.8	0.7	208315	1	AP003017	AP003017 Mesorhizobium
18	55.4	0.7	217594	1	AP002086	AP002086 Agrobacterium
19	54.6	0.6	5644	1	AF061070	AF061070 Pseudomonas
20	54.2	0.6	10137	1	AE009430	AE009430 Agrobacterium
21	54.2	0.6	14736	1	AE007938	AE007938 Agrobacterium
22	53	0.6	5142	1	SCF20	SCF20 Streptomyces
23	52.4	0.6	318	9	AF147369	AF147369 Homo sapi
24	52.4	0.6	349116	1	AP003003	AP003003 Mesorhizobium
25	52	0.6	12612	1	AE007883	AE007883 Agrobacterium
26	52	0.6	13029	1	AE008935	AE008935 Agrobacterium
27	52	0.6	71684	1	AF447814	AF447814 Escherichia
28	51.6	0.6	204050	1	AL646070	AL646070 Ralstonia
29	51.2	0.6	13457	1	AE007303	AE007303 Sinorhizobium
30	50.8	0.6	14466	1	AE007274	AE007274 Sinorhizobium
31	50.8	0.6	307551	1	AP003015	AP003015 Mesorhizobium
32	50.4	0.6	11383	1	AE004761	AE004761 Pseudomonas
33	49.8	0.6	11506	1	AE004793	AE004793 Pseudomonas
34	49.6	0.6	10441	1	AE009210	AE009210 Agrobacterium
35	49.6	0.6	11531	1	AE008176	AE008176 Agrobacterium
36	49.6	0.6	215050	1	AL646057	AL646057 Ralstonia
37	49.2	0.6	81	6	AX019311	AX019311 Sequence
38	49	0.6	338000	1	SMES91792	SMES91792 Sinorhizobium
39	48.8	0.6	349498	1	AP003002	AP003002 Mesorhizobium
40	48.6	0.6	3233	6	AX334910	AX334910 Sequence
41	48.6	0.6	3233	9	D50928	D50928 Human mRNA
42	48.6	0.6	194140	1	AF242881	AF242881 Agrobacterium
43	48.6	0.6	307551	1	AP003015	AP003015 Mesorhizobium
44	48.6	0.6	314150	1	CJ11168X1	CJ11168X1 Campylobacter
45	48.4	0.6	1987	1	PASSSDAF	X78478 P.aeruginos

## ALIGNMENTS

RESULT 1  
AX281574 LOCUS AX281574 8509 bp DNA  
DEFINITION Sequence 2 from Patent WO0177347.  
ACCESSION AX281574  
VERSION AX281574.1 GI:16608826  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
synthetic construct  
artificial sequence.  
REFERENCE  
1 (sites)  
AUTHORS D'Elia J.  
TITLE Ketogulonigenium shuttle vectors  
JOURNAL Patent: WO 0177347-A 2 18-OCT-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
source  
1. 8509  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="padm281"

BASE COUNT	ORIGIN
2065 a	2239 c
2345 g	1860 t



Query Match 100.0%; Score 8509; DB 6; Length 8509;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 8509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggtaacgcatlgtgcttgatglttggagaaaagcgaaagaccgagccagltgtgggt 60  
DB 1 TGGTAACGCATGTGGCTTGATGTTGAGAAAAGCGAAAAGACCCGCCACATGTGGGT 60

QY 61 agagcgtgatatagtgcgaagaccgtgactgttgcatacgaaccctgttactctgc 120  
DB 61 AGAGCGTGATATGTGCAAGACCTGATGCTTGACATCGAATCCGCTGTACTCTGC 120

QY 121 atcgcgctgatacagcgcgtcgcgaatgggaaagcccgctatgttcgtcagcagc 180  
DB 121 ATCGTCGCTGATACGCTGCTGCGGATGCGGAAAAGCCGCTTATGGTGTGACAGCAG 180

QY 181 ccttaaggcgaatgcggaacttggcccatgcccgaatcgtgtcgtttcaacatcgcgcat 240  
DB 181 CCTTAAGGCGATGCGCGACTTGGCCATGCGCATGTGGTGTTCACCATCGCGCGAT 240

QY 241 taacgaactggagatgatactctagaagcgtttaacgagaaatctggtttaacgcaataag 300  
DB 241 TACGCAACTGGAGATGATCTTAGACGGGTTAACGAGACATCTGTTAACGCCATAAAG 300

QY 301 ctgcgcatgaaatagagcgaaatctgcgttgcgcgccccgttctcagcgtgtct 360  
DB 301 CTGCGGCATGAAATAGGCGGACATCTGCGCTTGCGCCGCCCGCTTCACACCGGTCT 360

QY 361 gctctctgcgtcatgagcaagcagaaatcgcgttcaatacgggtctgatalccagacc 420  
DB 361 GCTCTGCTGCTCATGCGCGACGAGATCCGTTTACATGCGGCTGTGATTCAGACCC 420

QY 421 gccgccttgaacatgcgaacacatcggcatcaagccgcgcgagatctgttctgtat 480  
DB 421 GCCCGCTTGACCATGACGACATCGGCATCAAGCCGCGGATGATGCTGCTTGAT 480

QY 481 cgggagatagaagcggccacgcgtcgcgtctcccatlbggcttgcgtlaagctcgggaac 540  
DB 481 CGGGAGATGAAAGCGGCCACGCGCTGCGCTGCCATTTGGGCTGAGCTCGGGAAC 540

QY 541 atcgttgtgtcatctcgtcggggcggaagcgtccagcgcgcgaattcttcgcgcg 600  
DB 541 ATCGTTGGTGTGATCTGCTCGGGGCGGACGACGTCACCCGCGCAATTTTGGGCGG 600

QY 601 ctcctcgtgaagcggagcgcgtatcgaaggtgtatcaacatctggatctcttc 660  
DB 601 CTCCTCGGTAAGACGCGGACGCGTATCGAAGGTGATTAAACCATTTGGCATATCTTCC 660

QY 661 ttctctgcgtgttagcgcgcgcgcgaacgaatgtagcgcgactlctcgcgcgcgcgcgc 720  
DB 661 TTTCCTGCGGTGATGCGGCGCGCGCAATGATGCGGATCTGTCGACCGGATCGGGCG 720

QY 721 cagcctcgtatgatacaggtgagcaacagaaagcgcgcgcgcgcgcgcgcgcgcgcgc 780  
DB 721 CAGCCTCGATGATCAGGTGGGCAACGAGAAAGCGGACGCGCATGATCTGCGCAACG 780

QY 781 ttgcgaagcgtatctccgcgcctcgaatcctatcctcgaacgcgtcaggtgcaagcgtgcg 840  
DB 781 TTTCGACGCGGATTTCCCGCCCTCGATCTCTGAAACGCTGAGGTGCAAGGATCGG 840

QY 841 cgaacacatgacaaagcatctcgaacgcgaatgccaatgtctcttctcgttctcgcgc 900  
DB 841 CGAACAATGACACAGCATCTCGAACCAGATGCAATCTTTTGTGTTCCGCTT 900

QY 901 tggcggatccccagataaacgcgccttctatgycagaatataactaacatltgagtt 960  
DB 901 TGGCGGGATCCAGATAAACCGCATCTTATGCGAGAAATTAATACATTTGATGAT 960

QY 961 attcaatgcaagtgcgaggtcaaatcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1020  
DB 961 ATTCAATGCGCAAGTCCGAGTTCAATACGCCGCCCAACCGCAACTGATTTGTTCTAC 1020

QY 1021 tcaacgcgccttltgataagaagcttgcatgataaacaccgcgcgcgtctcacaacaaata 1080  
DB 1021 TCACCGCGGCTTTTAAATAGAAAGCTTGATGATTAACACCCGCGGCTCTCAACAAATA 1080

QY 1081 aggcaatccgcgcgcgtgcgcgaatctgcgccttgcgtatgcaagtcctgtgttcca 1140  
DB 1081 AGGCAATCCGCGCGGCTGCGCAATCTGCGCTTGTGTGATGCAAGGTCTGTGTTC 1140

QY 1141 tacttgagagatgtaaggaattgccccggaatgagacccaagcagacaccacaagcc 1200  
DB 1141 TACTGCAAGAGCATCAAGGAATTCGCCGATGAGCACACGACGACACCAACACGCC 1200

QY 1201 ggccttgaacaaaggcgcgtgtgtcgggaaaaaagcgcgcgcgtgacactgacagaattgc 1260  
DB 1201 GGCCTGGAACAAGGCGCGGCTTGTGCGGAAAAAGCGCGGCTGACACCTGACCAATGTC 1260

QY 1261 cctgatacgtctcatctcctgcgcgaagaaagcgtgtggcggaatctgcgtctgttaagct 1320  
DB 1261 CCTGATCCGTCATCTCCGCGCAGGAAGGCGGCGGATCTGCGCTGTGTTCAAGCT 1320

QY 1321 ggcgaatgaacacagtttgacgcgcgtcgaacctgtgcgcctgcgcgtctcgaatgtgc 1380  
DB 1321 GCGCATCGACACCAAGTTTTCGCGGCTCGGACCTGCTGCGCTGCGGCTGCGGATGTGC 1380

QY 1381 gacccacgtgtctgtcgtgaagatcgtcgaatccgcgaagaaagacccgagccgca 1440  
DB 1381 GACCCAGCTGTGTCGCGTGAAGATGTCGAGATCCGCGCAGAAAGACCGAGCGCGCA 1440

QY 1441 tgtcgcgcgcgttaacggccgcgtctgcgaagggaaacgcgcgcgcgcgcgcgcgcgcgc 1500  
DB 1441 TGTCCGCCCTGACAGGCGCGCTGCGGAGGGAACCGCGAGACCTCGCGGCTTATCT 1500

QY 1501 cgcgcgccttgaacagcgcgcgtcacagctgtgttcaacgcgaagcagcgcgcgtgtgc 1560  
DB 1501 CGCGGCTGTGAAGACGCGCTGACAGTGTGCTGTTACCGGACAGGCGATCGCGTGC 1560

QY 1561 ccaacacacacttaagcagaagccagctgtgcgcgtgtcaagctcgtcgtcgaagagc 1620  
DB 1561 CCACACCCACCTTAAAGGAGAGACGAGCTGTGGCGCTGTTCAAGTCTGCTGGAAGAGC 1620

QY 1621 ggcgcgtcgaatcccaagccttaagggctgacgcgcgtgcgcgcgcgcgcgcgcgcgcgc 1680  
DB 1621 GCGGCTGATCCAGCTTACGCGGCTGACGTGCGCTGCGCCCAACTTCCACGCCACT 1680

QY 1681 ctacgc 1740  
DB 1681 CTACGCGGAGACCGGCAATCTGCGCGCGCACAGTGTGCTGCGGCGATGCCAGCATCA 1740

QY 1741 gagaccaaagagttacatcgc 1800  
DB 1741 GAGCAACAAGAGTACATGCGGACGACGACCGGCGCTCGATATGCAAGGAGTA 1800

QY 1801 tcaacttaaccatgaggaacctatctgagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1860  
DB 1801 TCACTCTAACCCATGAGAACCTATCTCGAAGAGCGCATTCGCCGCAACACAGCAGC 1860

QY 1861 gttctacgc 1920  
DB 1861 GTTCTACCGCATGCGGCTGCTGCGCAACCTGTTGGGGAATGAGACGCTGATACGAGAA 1920

QY 1921 gggcgcgcgttgcatcagc 1980  
DB 1921 GGGCCCATTTGGCATAGGCGCGCATCTCGGCTCGATTTGTGAGAGTAACAAAGATTC 1980

QY 1981 catcgcgtcgaatgc 2040  
DB 1981 CATCGCTGCGATGCTGCCATGAGACCGCAAGCTCAGCGCGGATATGGCTCGAGGC 2040

QY 2041 catccagattgacatgttcccaagggcataaagccatcaaltgtaagagtgcaagcgcga 2100  
DB 2041 CATCCAGATTGACATGTCTCCAGGCGCATAACGCGCATCAATGTAAAGTGCACAAACGGA 2100

QY 2101 gcaagcacaagccattcaacagtgaggtgc 2160



Db 2101 GCAAGCAAAAGCCATTTCACAGTGGGTGGCAGATGTTCCGTTCACAGTGAAGCGCT 2160  
Qy 2161 gatgctgttcccaagccacagagctgatagcacaaagcaacgaggtctgcgccacagac 2220  
Db 2161 GATGCTGTTCACAGCCACAGACTGATAGCACAAAGCAAGGGGTGTGGCCACAGAC 2220  
Qy 2221 cggcttcgcgcgcaaccccgcaagaaacgaggttaaaatgagcatttcgcgcaaaaacgctg 2280  
Db 2221 CGGTTCCGCCGCCACCCCGCAGAAACGAGTAAATGGCGATTTCGCCCAAAAAACCGCTG 2280  
Qy 2281 caaatgtagtgcgaatcaccalccagtttcaacctgtaaaaccgctgcctcaaatgaaagag 2340  
Db 2281 CAATGATGCGCAATACACATCCAGTTTCATCTCGAACCCTGTCGTCACATGAACGAG 2340  
Qy 2341 caagccatactccaagcccccaagaaacgaggtgcgagcactacagatgagcagtgctcgtg 2400  
Db 2341 CAGGCCATCTCATCCAAAGCCCGAAGAACGGGTGGCGACATACATGAGCGATGTTCTGG 2400  
Qy 2401 ctcataggtctgcaagagccctgcgaacagtgatttcaacgttagatctgcaggtctcttggc 2460  
Db 2401 CTGATAGGCTGCAAGGCCCTTGCAACAGTGTTCACCGTGAATTCGACAGGCTCTTTGGC 2460  
Qy 2461 tctcccgcaagagccacactcaaggttgagcagctagcgcctctaaggttcaacagtgaaatcg 2520  
Db 2461 TCTCCCGCAAGAGCCACTCAGGGGTGAGCAGCTAGCGCTCTAGTTACACAGTGAATCG 2520  
Qy 2521 ctgagagagcgtgcgaggtcttattggttgcgtgcgttcacgttgcacatcgccgacatgcacat 2580  
Db 2521 CTGAGGAGCGCTTGCGGGGCTTATGGTTGGCTGGTGCACGTTGGCCATCGGAGTGGACAT 2580  
Qy 2581 acgatgctctctaagcagctcgaatcctgaagctcacaagtcggaanaatacagctcacaanaa 2640  
Db 2581 ACGATGGCTTCTAGCGCAGTGAATCCTGAGGCTTCAGGTGGAAAAATACGCTCCAAAAA 2640  
Qy 2641 agcccttgacccaatcttgganaaaatctgcttgaanaagttcgttctaanaaaacttggaaag 2700  
Db 2641 AGCCCTACCAAAATCTTGGAAAAATTTGCTTGAANAAGTTTGCCTTCTAANAACCTGGAAAG 2700  
Qy 2701 agatattgcaagagatccctctacagatgctgtagaagtaagcagtggaacaaacgacat 2760  
Db 2701 AGATATGCACAGAGATCCCTTACGAGTGTGAGAGTAATGCACTGGACAAACAAAGCGCAT 2760  
Qy 2761 tttttgccccagtagagaataagtagtgcgtatttcttggagagaatttgcctctaagtaga 2820  
Db 2761 TTTTGGCCCAAGTAGAGGTATGAGTGGTATTTTGGGAGATTTTGCTTCAGTAGGA 2820  
Qy 2821 gtaacgcttgcgttaaaatttgcgtgattgcgttcgaatcgaatcgaacacgagctgcgttg 2880  
Db 2821 GTAACGGGTGGTTAAATTTGGCTTGATTTGGCGGTTCAAAATCGACACGAGCTGCCGTG 2880  
Qy 2881 gtcgtattgcgactgcgccgcaatttggcaacttgcagggccatcccccctgaactctgcg 2940  
Db 2881 GTCGTATTTCATCTGCCCCCAATTTGGGCACTTGGCAGGCCATTCGCCCTGAACCTTCTGGCG 2940  
Qy 2941 atgacccatttcgaagcaatgcgtgcgaatcattagaaatttgcgtgagtgagtcgag 3000  
Db 2941 ATGACCTTTTCGAAGCAATGCGTGCGAATTCATAGAAATTTTGCTGAGGCTTAGCGG 3000  
Qy 3001 cctctgacagaggtgcgtgcgagagatcctgcgttcaggttagagggcgacaatgagagagtg 3060  
Db 3001 CTCGGAACGAGGCTGCTCGGGAGATCTGCTCAGGTGAGGGCGACAATGAGAGAGTG 3060  
Qy 3061 tttagttcccccgtatcgtctctctgcgtgcgtatctgggtgcacactgcgcccgagacatg 3120  
Db 3061 TTATTTCCCTCTTATCTGCTCTCTGCTGGCGCATTTGGGTCATCTGCGCGACATATG 3120  
Qy 3121 atattcgcctagaagatctactgatagttttcgcctgcgtgcgagctgcgtgcgtgcgagc 3180  
Db 3121 ATATTCCGCTAGAGAGATTACTGATAGTTTCTGCTGTCGGGCTTGCGGCTTGTGCGGGC 3180  
Qy 3181 ttgtcggagctgtgcggagcgtgtccctctgtccgcgctgtcctaactttcacaaatcaa 3240  
|||||

Db 3181 TTGTGCGGCTTGTCGGGCTGTCCCTCTGTGTGCCGCGCTGTCCATCTTTTCACATCA 3240  
Qy 3241 aaatggcgagagccctctgttctatagtttctatagtttctatagtttctatagtttctatag 3300  
Db 3241 AAAATGGCGAAGCCCTTCTTGTCTATAGTTCTTATAGTTCTATAGTTCTATAGTTCTATAG 3300  
Qy 3301 atttcaatagcttattcgccttaaaaggaatlaattggcgccgcaaaaaggagtagtgg 3360  
Db 3301 ATTATCAATACCTATTTCGCTTAAAGGAGTAAATTTGGCCGCCGCAAAAGGAGTAAATTTGG 3360  
Qy 3361 ccgcaaaaggagtagtggcgcaaaaggagtagtggcgatalcggttgcgttctac 3420  
Db 3361 CCGCAAAAGGAGTAAATTTGGCCGCCGCAAAAGGAGTAAATTTGGCGATATCGGTTGTTAC 3420  
Qy 3421 atgggaggaatcccttataatcttctccccaagggaagaacaaacaaatgagtcgcgag 3480  
Db 3421 ATGGGAGGAATCCCTTAAATCATTTTCTCCCATAGGGAACAAACAACTATGACCCGCGAG 3480  
Qy 3481 accgggctctgcagacagacaanaaacgtgctccctgcgaggttgcgagaggttctata 3540  
Db 3481 ACCGGCCTTCGACAGACAAACAAACTGTGCTCCTGCGAGGTGGCGAGAGGGGTCTATA 3540  
Qy 3541 tgcgaatlcgccccgcctgcgagcgtcgaagctcaatgataltaatgataagcactgcg 3600  
Db 3541 TTGCGCAATCCCGCCCGCTGCGAGCGCTCAAGCTCATGATTAATGATAGCCACTGCGG 3600  
Qy 3601 gcggcgcgcatggtcgtatgattgcgcgaatgaatgagcgttgcgcggaattgcgcaatcg 3660  
Db 3601 TGCGCCGCTATGCTGATGATGTGCGCATGAAATGCGGCTGCGCACATTCGCGCAATCG 3660  
Qy 3661 accgcatgaanaaacatlgagacgttgagacgtgcgaacccgctcttcgagagctgaacgctg 3720  
Db 3661 ACGGCAATGAANAACCATATGACCGCTGAGACCTGACCCCGCTTTCGAGAGAGCTAGCCGCTG 3720  
Qy 3721 cgtgtgtgacccatgatagccttcgaaagatgactgcgtgagcgtgcgagcttgcgtg 3780  
Db 3721 CGGTGTTGACCCCATGATGATGACCTGCAAAAGATGATGCTGACAGTGGCGGCTTGGTGCATG 3780  
Qy 3781 aggcgcgaatagactacccgcgaggaagcaagcgcggaactcctctgtagcgtgcgagactcc 3840  
Db 3781 AGGCGCAATGAGCTACCGCCGAGAGGCAAGCGGGAACCTCTGATGACGAGGACCTTCC 3840  
Qy 3841 ggaatatactcgcgtcgtatgc 3900  
Db 3841 GGAGTACATTCGCTGCTGATGAGCGGCGGAGTGCMAACCACTTGCGCCATTTCCACGCTCAA 3900  
Qy 3901 cgtatccatctcgttagtagtattcgcgtgcgttgcgttccagcaagctcctagtcgcg 3960  
Db 3901 CGGTATTTCATCTCGGTAGTAAGTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3960  
Qy 3961 ccaatcttgcgtatgcgagcgcggaacacttaccgttcccgagtttgcggcgcttcccttg 4020  
Db 3961 CCAATCTTGTGATGAGTGAAGCGCGCAAAACCTTACGCTCCGAGTGGGGCGCTCTCTTGG 4020  
Qy 4021 gaagtcccgaggaagaatgctgcgtggaagcagcttaacagatttctctcaaacctg 4080  
Db 4021 GAGTGCCCGAGAGAAAGATGCTTCTTGGAACGACGTTAAACAGATTTCTCTCAAACTG 4080  
Qy 4081 caactgtagatgaacacacattatcgcgtcgaatlgacggcgcaagcgacgaactg 4140  
Db 4081 CACTGTGATGAGATCAACCATTTATTCGCTGTGACATTTGACGGCAAAAGCGCACGAAGATTG 4140  
Qy 4141 ggcgtaggtgcgaaggtgaactatagcttgcgtggaagtgaaagacgaacccaacgctgcgca 4200  
Db 4141 GCCGTAGCGTGCAAGTGTGATAGGCTGGGAAGTAAAGACGACCAACCGTCGCA 4200  
Qy 4201 ggcgcaagctggcggttcccaagtcggttcgagatgcctgcgttcgcaagggcgagcgagaa 4260  
Db 4201 GCGCGAGACTTGCGGGTTCCAAAGTTCGCTCAGATGCTGCTGCGAGAGGGCGAGGAGAA 4260  
Qy 4261 cgatagccctccttcccaagcgagcgagatcaactacagtcacgcttgcgtgagc 4320  
Db 4261 CGATAGCCCTCTCCCAAGAGCGGGGAGATCACTTACAGTCCACGTTGGCTGAGAC 4320



OY	4321	tgaacgtctcgtcgcgcgaacaaagaaacga	ctgtatcgtccccaactccgcgcgtt	4380	
Db	4321	TGAACGCTCTGCTGGCAGCAACAAAGACAA	MCATCTGATCCCTCTAGACTTCGGCGCTT	4380	
OY	4381	tcctgcggagaaagagcgctgcctcgtgaac	ctcgaacctcgaacacgcgaanaactcgtttta	4440	
Db	4381	TCTGTCGGAGAAAGCGCTGCTGTGGAGCTT	CGAACATCGAAAACGTTTATTAGATT	4440	
OY	4441	tcctgcgaagtagaggaggtttgagttttg	agttatccaccgaatagtgcttaaatg	4500	
Db	4441	TCTGCGCAAGGTAGCGGAGGTTTGACTTT	GAGATTTCACCGCAATATAGTTTAAATG	4500	
OY	4501	acttcctcgaacacga	tgtgcataatagcgtlaagactatgaat	latacgcgctcggaagcg	4560
Db	4501	ACTTCCTGTAACATGTGCAATATAGCGTAA	AGACTATATAATATACACGCTGCAAGGC	4560	
OY	4561	tgcgaagaacacggtgtg	gcgcgcgaacacataccgcgcgcctaaagc	gcgtlaaat	4620
Db	4561	TGCAAAAGCAACGGGtgtgcgcacccgCA	ACCATACTTCGGCGCTATAAAAGCGGTAAAT	4620	
OY	4621	ttccgcgtaaagaagatg	aatctcggcga	ctggttatagatctcgcgaatctgcagagt	4680
Db	4621	TTCGCGTAAAAAGATGATCTGTGGGCA	TGTGGGTATTAGTCTCTCACAATTTGCACAGT	4680	
OY	4681	gtttccctccattccaagaata	lataccggaacacctaaacgcgaatata	tgtlaagcg	4740
Db	4681	GTTTCCCTCCATTTCAAAAGAAATATAC	CCGCAACCTTATACACGAAGTATGTGTAAGCG	4740	
OY	4741	tgatgaacacatgaaatga	atgaacctcagaaatcagcgcatlagagcg	tgaagttcgacttt	4800
Db	4741	TGATGAACACATGAAATGATCACTCA	GAATATCAGCGCATTTAGCGCTGTAAGTTCGACTTT	4800	
OY	4801	acgcgaatgctttatctg	atgcgcgaagagatcgcgacaat	ctgcgcgcgaatgcgcgcgcg	4860
Db	4801	ACGGGATGCTTTATCTATGTCCAGGAGAT	TCCGCACAATATGGCCGCATATGGCCGAGCG	4860	
OY	4861	tccttcaatcttcataccacga	atlgagagagaacgcgccccctc	taaaaaaagaatgtg	4920
Db	4861	TCTTTCATTTTCATCACCGATGAGAGAG	AGAACCCGCCCTCCATAAAAACAAAGATTG	GTG	4920
OY	4921	gaagatattccgatccctgcgcgtcttaaga	agactgcgtctttaaactgaatcaagctcta	4980	
Db	4921	GAAGATATTCTGATCTCGGTTTCAGAG	CGCTTGCCCTTAAAACTGAAATCAGCATTTCA	4980	
OY	4981	gcgatgctgataagaagta	aatatataagccaataatagacgcgcac	tttccatccaatata	5040
Db	4981	GGATGCTGTATGAAGATTAATATATG	CCACATATGAGCGGCATTTTCAATTCACATACA	5040	
OY	5041	gcctcactatgataatata	aatgatgatataatccaat	ctgagagaagaatcttaacta	5100
Db	5041	GCTCATCATGTGATCATATATCAATG	ATATCATATCAATGGAAGAAATTTACATATGA	5100	
OY	5101	tcaacagatcatcaacga	catgtgtttgtgttttcttaagtcgaata	taactatcgcgtg	5160
Db	5101	TCACAGGATCATCACGACTTGTGTGT	TTTGTATTTTCTAAAGTCTTAACATATCTATCGCTGG	5160	
OY	5161	cctctaagaagatctg	actatctgcagta	lctcaaccttgggtttcaagacgaataatag	5220
Db	5161	CCCTAAAGAAGATTGTATGATTTGCAG	TATCTCATCTCCTTGCTGTTTCAGACCGTAATTTACAG	5220	
OY	5221	cttgttcgaagccggttttttttggcag	agaagacagtttttga	tgtgtgtaataaactg	5280
Db	5221	CTTGTGTGAAGCCGGTTTTTTTTTCC	AGAAGACACGTTTTTGTATGTGTAAATAACTG	5280	
OY	5281	ctagatctc	atgatgtgtaata	latacacgcatcgcgggcacaacactctta	5340
Db	5281	CTACGATATCTATGATGTTAACTATACA	AGCAATCATCCGGGCAACACACCTCTTATGGA	5340	
OY	5341	agatgataata	tgcccgagaagctatgcgcggaanaacgcga	latlgagtgcaagcccg	5400
Db	5341	AGATGATATTTATGCCCCAGAAAGTACT	GGCGGAAAAACGATATTGAAGCGCACAGCCCG	5400	

QY	5401	cacttttagcggaagccctataacagatgacccaacaaacatagaagccagattttagggaaaa	5460
Db	5401	CACTTTAGAGCGGAAACCTATATACAGATACCAAAACACTAGAAAGCCAGATTGGGAAA	5460
QY	5461	taaggaaatgctccctcgaagcgctctgagctagcttagcagctcctttaaataatgttag	5520
Db	5461	TAGGGAAAGTCCTCCAGAGCGCTGGGGCTAGCTAACAGATCTTTTATTAATAGGTAG	5520
QY	5521	tacagaagaacagaacattatacagggccgaaccccgacgagatcctgttgaatctgt	5580
Db	5521	TACAGAAAGACAGCAATTTATACAGGCGCGACATCCGAGGGAGATCTGTTGTAATCTGT	5580
QY	5581	acccgatgacatctcagaatctaaacaagaatcgagagccggaaccgagctgtcttccagc	5640
Db	5581	ACCCGATGACATCTCAGATCTTAACAAATCGAGAGCCGGAAACGGCTGTGTTCCAGC	5640
QY	5641	atcagttctgtttagagagatgtgagacgagctgacagatttgcacaccggttaccctgccc	5700
Db	5641	ATCAGTTCTGTTTAGAGAGATGTGAGACGCGCTGACAGATTCACATCCCGATCTAGTCC	5700
QY	5701	atcagctcgaatgacgaagcaactagctgtggttaacagatgagaaactaaagattacagttatga	5760
Db	5701	ATCAGCTCGAATGAGGAGCAACTAGCTGGGTAAACGATGGAATGAACATAACGTTATGTA	5760
QY	5761	agaaaagaagaggttggctagaaatcggagagctacagctctgcacatgtgcataatggagt	5820
Db	5761	AGAAAAGAGCGTGGCTTAAGAAATCGAGACCTAGCTGCAATGCAATGCAATTAATGGAAT	5820
QY	5821	aagtgcgcgctgcgatccagctgtaactctctctggaatcccaaccatgtgatctgtaattgg	5880
Db	5821	AAGTGCAGCGGTGATTCAGGTGAATCTTCCGCAATCCCAACATGCTATCGTTAATGG	5880
QY	5881	gcaattgcacccctggagttcttcgagatattcagatcccaagagccgaagatcccat	5940
Db	5881	GCAATTGCACCCCTGGAGTTTTCGAGATATCTTACGATTCACAAAGCCAGAAAGCTCCAT	5940
QY	5941	atccaccaagaatgtctgaataatgtgggctcgtgagaaaccacagataataccgatctatc	6000
Db	5941	ATCCACCCAAAGTGTGAATAATGTGGGCTCGAGAACTCAAGATATATACCTATCTATTC	6000
QY	6001	tatgtaagcttcgcagctgcgcctcogaatgatcaacagatagatgtatgcaataacatga	6060
Db	6001	TAGTGAAGTTCGCACTGCCGCTCGCAAAATGATCAACGATAGATGATCAATTAATCTGA	6060
QY	6061	ttccsagaaatttagctgtgtatattcttcctgtgaagaagagacaggaattactatcccta	6120
Db	6061	TTTCAGAGATTTAGCTGTGTTATATCTTCTCTGAAAGAGACAGATTACTACTTCACCTA	6120
QY	6121	tctgtgcggaattccaacctcaaaaacagatgtgatttgaatgtctcagaagaggtgaaatcac	6180
Db	6121	TTCTGGCGGATTTCAACCTCAAAAACAGATGATTTTCAATAGTCAGGACAGGTGAATCAC	6180
QY	6181	ccgctgaatttccaccaaggtgagctcctgttagatccaagctccaagaagatlaaacgltta	6240
Db	6181	CCCGTGAATTTCCACCAAGGTGAGCTCTGTGATCAACACTCTCAAGAGATTAACGTTTAA	6240
QY	6241	atccatcccatagatacaatacagatagatgtgtcagttccggtcagatctctgtgtatccggt	6300
Db	6241	ATCCATCTCATAGATCAACATAGATAGGTGTCAAGTCCCGGATCTGTGTGATCCGCT	6300
QY	6301	ttagatgaatctgtccgagctcttaacatacccccggtttaaaccggtcttatacaagag	6360
Db	6301	TTAGATGAATCTGTCCGGCTCTTAGCATATCCCGCGGTAAACCGTCTTTACAAGAG	6360
QY	6361	aaatagagcgagctcgaagcgctcttagcgcgatgtcccggtccaaacggagatgtgtagagc	6420
Db	6361	AAAGTCAGCGGCTCGAAGCGGCTTAGAGCGATGCCCGGCGCCAAACGGATGTGATAGC	6420
QY	6421	gaacaagaagaagcgctagctatgtgtctgtgcccgtcccgagagctgcaagccgagaatcc	6480
Db	6421	GAACAAGCAAAAGCCCTAGCTATGTGCTGTGCCCGTCCGGAAAGTCGACCGCGAGATCC	6480
QY	6481	ggaaaaaagaataatactatgagcgagatattttagatagaacacatcctttctattia	6540



Db	6481	GGAAAAAGAAAAATACATGGCAGGGATTATTGGATAGGACACAACTCTTTATTA	6540
Oy	6541	atatcaacaagaatatyggatcgcgcgcgcgtgtaaccattcgatacaatccaaatcc	6600
Db	6541	ATATACAAACAAGATGTGGCATGCGCCGCGCGGTGATCCTCATTCGATCAATCCAAATCC	6600
Oy	6601	tgaagctgactatgcccctacgcacgcgcgaacatcgtgtccgtaattgaigaac	6660
Db	6601	TGAAGCTGACTATGCCCTACGCATGCGCACCATGTGGTCCGTCATTGATGATCGTGAAC	6660
Oy	6661	cgagccactcgtcgtcccgaaatccaagccctatgcttgctgcgcggaagtcgt	6720
Db	6661	CGCACCTACCTGCTGCTCCCGCAATCCAAAGCCCTTATGTGTGGTCTGCGGACAGGTCCTT	6720
Oy	6721	gcgctgttcgaactcctcgtttgaagggctatccgatacgaagctcatgtctcgcgcgtg	6780
Db	6721	GCCTGTTCCACTCTCTCTGTTGAAGGGCTATCCGATGACGACTTATGTCTGGGGGTG	6780
Oy	6781	gaagaagagaccaaagcagaagctgcgaatgctacaatltcatcgaagaatlcatcgcgcga	6840
Db	6781	GACGAGAGACCAAGGACGAGAGCTGCATGCTCAATTCATCATGAGAAATTATGGCCGGA	6840
Oy	6841	atgatgaacgagccgactagtcgcgaagcgcgcgcagtcgcctcgtcgtcgaagacag	6900
Db	6841	ATGATGAACGCGCCGACTATGTGGGAGCGGGCGGACGTGTCTGTCTGTCACGACAG	6900
Oy	6901	cagaagatgacctactcgtttgatccgctctgcgcgcgaatctctcgaagaagccaaac	6960
Db	6901	CAGCGGATGACTCTACTCTGTTGATCGCTTGCGGGACATTTCTGTGAGAAACCAACAC	6960
Oy	6961	gcgcgcacaacgacgcgcgcgcgtgcgtgcgcgaacaaacgtatatactagacctgtcgcg	7020
Db	6961	GCGCGCAACAGCAACGCGGGGGGTGGTGGCAAAAAGCCTATATCTAGACTCGTGTGG	7020
Oy	7021	gaccgcgaatccgaagaacatccgatagaagcaagcaatgagltcgaatcaactaagtt	7080
Db	7021	GACCGGATCCGAAGAACTCCGATGAAGACGAAGCAATGATTCGGAATCACTTACGCT	7080
Oy	7081	ctctcttctcaatgaagccgcgcgcgcgaagcagaagcaacatggttaagaatgga	7140
Db	7081	CTCTCTTCTCAATGAACCCGCCCGACAGCGCACAGSCACCCACTGTTCAAGGTGGGA	7140
Oy	7141	tcgaatactggaatltccatacagaacgcacgtgaagggltbatgcgaagtgaagacc	7200
Db	7141	TGCAATACGATATTCCTTACAGACGACACTTGAGGGTTGATTTGCCAAGGTGAAGAC	7200
Oy	7201	gaattcaatagtgatcgcgatctgggaaagggcgtgcgcgaagaacccctgcgcgg	7260
Db	7201	GAAATTCATCATGTGTATGATGATGGGAAAGAGGGCTGGCGGAAGACACCTCGCCGG	7260
Oy	7261	ctgacccgcgcatactctgaaagacgcgcgataacttlttaactgaagcgcgcgcgcgcg	7320
Db	7261	TTTGACCGCGTCACTCTGGAAGACGAGGGGATCAACTTTTTCATCTGAACGCACCACTGC	7320
Oy	7321	gttbatcgcggtgcgcgcacatctcgtgcgcgcgaatgcacgcgcgcgcgaagaatgtcgaag	7380
Db	7321	GTTGATGGGGTGCAGACATCTTCTGTGGGCCCAATGAGAGGGGGCAGCAAACTGTCCAG	7380
Oy	7381	gcagaacctgcatgltcatgatcaagtcacaaatggtccagcgcgatcgcgcgcgcgaagaa	7440
Db	7381	GCAGACCTGCTGATGTGATGATCACGTCAAAATGCTCCACGCGGATCGGCCGCGAGGAA	7440
Oy	7441	atcggcgcgtttgtcgcgcacataaacaagctcctcgtgtgcgcgcgaacaagataagtcgc	7500
Db	7441	ATCGCGCGCTTGTGTGAGACATTAACAAAGGCTTGTGGTGGCCCAACAAAGTCAAGTGTGC	7500
Oy	7501	gactcgcgtccgaagggcctcgtcgtcgtcgtcgtcgtatgatagtcgtcgtcgtcgtcgtcgt	7560
Db	7501	GATCTGGTCCGAAGGCCCTGTCTGTGCTGTGCGGATTAATGATGTGCTATATATCTACGG	7560
Oy	7561	aacttacaagcgcgcgtcatcgcgcagatcgaagaacatggaatcgtatacgaagcga	7620

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AX281868	Sequence 1 from Patent WO0177159.	AX281868	AX281868.1	GI:16609119	unidentified.	unclassified.
Db	7561	AACTTTACAACGAGGAGTATCGGACAGGATTCGAAGCAACTGGGATTCGATATACACAGGCA	7620			
Qy	7621	ttcgaagaacagcttcacgcctctcgtgaacagagcatgcatcacacggggataacctcgcgcct	7680			
Db	7621	TTTCGAGAACACGTTCCGCTCGCTGTGAACAAGGATGGCATACCGGGGATTAACCTCGGCTCT	7680			
Qy	7681	ttgaacgcggtgctgctctcgtgctctatatatactcaacaacgcggatlltcgatlctccga	7740			
Db	7681	TTGAACGGGGGATGCTGCTCTGCTCATATATATCTCAACAACAGCGCCGGATTGGATTCCGA	7740			
Qy	7741	gagatgagcagaggttcgagcggggtcaatgcaggttcctcaagacgtctggttgtagaacgc	7800			
Db	7741	GGATGAGACGAGTTTCGACGGGGTCATATCCAGCTTCATGACCTCTGTGGTGGTAGAACAGC	7800			
Qy	7801	ctgcgtgctcaagcgcctctggttggccatctcgatcaagacacatcacaccacgcgcgcatacg	7860			
Db	7801	CTGCTGGTCAAGGGCTTCTGGTTGGCCATTTCGGATCGATGACACATACACCGCGCGCATATACG	7860			
Qy	7861	atcccgagatcacactgctgtgtaggccgcgaattcccaagtaagaagctgttgcgatgcatg	7920			
Db	7861	ATCCCGAGATCACCTGCGTGTAGGCGCGCATTTCCAGTACGAAACCTTTGGATGCCATG	7920			
Qy	7921	ggcgaagggggagagcgtctatcaggttgatagagcttaccatcgaagaagtgtcgtgaatg	7980			
Db	7921	GCGAAGGGGGGAGCGCTATCTCAGGTGGATGAGCGCTACCATCGAAGAAATTCCTGGAAATG	7980			
Qy	7981	caatatgcaaacaccccgagacctgtgtgcgtctgcgtccctatcagaagatcgtaactgg	8040			
Db	7981	CAATATGGCAAGCCCCGGACCTTCTGTGGCGTGTGCTGCTCTATCAGGGCATCGACTGG	8040			
Qy	8041	aacgagatcgacctgagcatgctcgaatcatactatctcccaagcggagcgtcacaagaaatgtg	8100			
Db	8041	AACGATCGACCTGGCATATGTCATATATCATATCCCAAGCGGAGCGTCAGAAAAATGTG	8100			
Qy	8101	ctgcgcgggagcgaatctcgtcccgagatcgcgcatctcaagaatcttggcgcggttaacagt	8160			
Db	8101	CTGCCCGGGCGCAATCTGCCCGAGCATGCGCATTCAGGAAATCTTGGCGCGGTTAAACAGT	8160			
Qy	8161	ttgggcaaacctgcgaacttttgcggagagatagaatatcgagaaagtgctgtgcatactc	8220			
Db	8161	TTGGGCAACCTGCAACTTTTGCGCGGAGATGGAATATCGAAGAAAGTGCGCTGCATATTC	8220			
Qy	8221	aggtcatgatattacccgaagcggcgctgtgatcttctacgaagcagcatatgatcccgcgctac	8280			
Db	8221	AGGTCATGGATTACCGGAGCGGCGTGTGATTTCTACGAGCAGCATATGATCCCGGCGAC	8280			
Qy	8281	cttgaactgtcgatgtaatactgcatctgcgcgaagtcgtgtgcgcaacgggaaagtgatc	8340			
Db	8281	CTTGAACGTGTGGAGTACTGTCATCTGCCCGAGTTCTGTGCGGAACGGGAAAGGTGTATTC	8340			
Qy	8341	cgggcgcctttgatagaattgttgcgagacagacgcgcatagaatgaagtgctgtctgtac	8400			
Db	8341	CGGCGCCGCTTTGATGAGAGTTGGTCCGAGCACACACCGCATGATAGAGTGCTTGTGTAC	8400			
Qy	8401	gcgaagagctgcgtcaatcttgcgcgaacctgtcgaagaaagcgctgtcgaacacctg	8460			
Db	8401	CGGAGAGAGCTGCGTCATCTTGTCTCGACACTGCTTGAAGAAAGCGCCTGTGCAACACCTTG	8460			
Qy	8461	cgggaacaccaaggaagctcgccgcgcgctatgttgtgtgcaaccggaaga	8509			
Db	8461	CGGGACACCAAGGCAAGCTCGCGCGCGCTATGTTGTGCACACGCGGACGA	8509			
RESULT	2					
LOCUS	AX281868	8509 bp	DNA	linear		PAT 02-NOV-2001
DEFINITION	Sequence 1 from Patent WO0177159.					
ACCESSION	AX281868					
VERSION	AX281868.1					
KEYWORDS	unidentified.					
SOURCE	unclassified.					
ORGANISM	unclassified.					







[illegible]

QY	3001	ctcgcgaagggtgtcgtcgcgcgagatctctgltctcaagtaaggcgacaatgagaggtg	3060
Db	3001	CTCTGACAGAGGGGTGCTGCGCGGAGAACTCTGTGCTCTAGTAAAGGCGACAAATGAGAGGTG	3060
QY	3061	ttaatgtccccctgtatcgtctctctgcgtgagcgatgtgtaacctgtccggacaatg	3120
Db	3061	TTAGTGTCCCCCTGTATCGCTCTCTCTCGTGGCGCATGTGGGTATCTCTCCGGACATATG	3120
QY	3121	aatacgcctagagagatlaactgaatgtctctgcgtctgcgtctgtcggcgtgtgtcggc	3180
Db	3121	ATATTTCCGCTAGAGAGATTACTGATNAGTTTCTGCTGTGCGGCTTGTGCGGCTGTGCGGCG	3180
QY	3181	ltgtcggcgtgtgtcggcgtgtccctctgtcccgctgtccctaacttttcaacaatcaa	3240
Db	3181	TTGTGCGGGCTGTGTGCGGCGCTGTCCCTTGTGTCGGCGCTGTCTCACTTTTTCACAAATCAA	3240
QY	3241	aaaatggtcggaagccctctgtctctatagtctctatagtctcatcgnaattacaata	3300
Db	3241	AAAAATGGCGGAAGCCCTTGTCTTGTCTATAGTCTTATATATTATATGAAATTCACACATA	3300
QY	3301	attatcaatagcttatcgtctttaaaggagtaattgtggcgcgaagaaggagtaattgtgg	3360
Db	3301	ATTATCAATNACTTATTTCGCTTTAAAGGAGATTAATTGGGCGCCAAAGAGGATTAATGGG	3360
QY	3361	ccgcgaagaaggagtaattgtggcgcgaagaaggagtaattgtggcgcgatalcgtgtgtttac	3420
Db	3361	CCGCAAAAAGGAGTAAATTGGGCGCGCAAAAAGGAGTAAATTGGGCGCATTAATCGGTTGTATAC	3420
QY	3421	atgtggagaggaatcccccttaactatctctcccaatgggaagacaaacaagtgtgcgcag	3480
Db	3421	ATGGGAGAGTAATCCCTTAATCAATTTCTCCCATGGGAAAGCAACACAAAGTGGCGCGAG	3480
QY	3481	acggcgcccttgacacaacaacaacatgtctccctctcgagatgtggcgaagggttacta	3540
Db	3481	ACCGGGCCTTGACACACAAACAAACTGTCTCTCCCTCCAGGTGGCGAGAGGGGCTATA	3540
QY	3541	ltgcgaatccgccccgcgtcgagcgctcaagctcatgtacattaaatgaatagccaactgcg	3600
Db	3541	TGCGCAATCCGCCCCCGCTCGAGGCGCTCAAGCTCATGCAATTAAATGATAGCACTGGCGG	3600
QY	3601	gctggccgaatgtcgtcatgtatgtgtgcgaatgtggctgtggcgaaattcgcgaaatcg	3660
Db	3601	GCGGCCGATGTCGATGATGTGCGCATTAATGTGGGTGGCGCAATTCGCGCAATTCG	3660
QY	3661	acggcaatgaanaaacatgacgtcggtgagagccgtgacccgcgtgtcttgaggagcttagccgtg	3720
Db	3661	ACGGCATGAANAACCATGACCGTGAAGCTTGAACCCCGCTGTTCACAGAGCTAGCCGCTG	3720
QY	3721	cggtgtctgaaccatgatacgtccctgtgcaagaatgtatctgtgacgtctggcgagctgtgtcgtat	3780
Db	3721	CGGTGTTCACCCATGATGACCTGTGCAAAAGTATGTGTGACAGTCCGGCGGCTTGGTGCATG	3780
QY	3781	aggcgcgaaatagactacgcgccaggaaggcaagcgcggaactcctagtgtgacgtgtgac	3840
Db	3781	AGGCGCGAATAGACTATACCGCCAGAGGCAAGCGCGCAACTCTTAATGTAGCACTGTGACCTTCC	3840
QY	3841	ggaatgaacttccgtcgtatgtggcgcggaatgcgaacaacatcgggcgactcttcgaacgttcaaa	3900
Db	3841	GGAATGAATTCGCTGATGATGGCGGAGTGCAGAACCACTGTGGCCATTCTCGACGCTCAAA	3900
QY	3901	cggtatccatcctcgtgtagtaaatlatctcgtgtgtctgttccagaacgtlctctatgtctcg	3960
Db	3901	CGGTATTCATCTCCTGATGATGAAGTAATTCGCTGTCTGTTCAGACACGTCCTTAATCTCG	3960
QY	3961	ccaatctgtatcgtgatagtagcgcgaaaaacttcaagtlcccgagttgtggcgctctctgt	4020
Db	3961	CCAATCTTGATCTCGATGATGCGCGCAAAAACCTTTACGCTCCCGAGTGTGGGGGCGCTCTTGG	4020
QY	4021	gagtgccccgaagggaataatgtgtcgttttgagaacagctttaaagatgttctctcaaaactg	4080
Db	4021	GAGTGCCCCGAGGGAATAATGTGTCTTGTGGACACGCTTAAACGATTTGTCTCAACACTG	4080
QY	4081	cactgtgataataacaacttlatcgtcgtctgtgacatttcgacgtgcgaagccgaaccgaagtgtg	4140







OY	6301	ttagatgaatcctgctccgctcttgacatacccccgctgbaaacctgtcttcaaaag	6360
Db	6301	TTAGATGAATCTGTCCGCGCTTGTGACATACCCCGCGGAACCCGTCTTTACAAGAG	6360
OY	6361	aaagtcaagcgctctgaagcgcgtcttaagccgattgcccggtgcccaacggtatgagtgaagc	6420
Db	6361	AAAGTCAGCGCGCTCGAAGCGGCTCTTAGCGCATGCCGCGGCCCAACGGGATGAGTAGAGC	6420
OY	6421	gaacccaagcgcccgactgatgtgtctcccgctcccggaagctctgaacgcgcaagatcc	6480
Db	6421	GAAACACCAAGCCCTCAGTATAGGCTCTGCCCTCCGGAAGCTCGACCGCGAGAATCC	6480
OY	6481	ggaaaaaagaataatacatctgacgacgatatcttgatagagacaaatcccttctatla	6540
Db	6481	GGAAAAAAGAAAAATACATGGCAGGAGATTATTGGATAGACACAACTCTTTCTATTAA	6540
OY	6541	atatacaacaagatatyggcattgcgcgcgctgatactcatatcgaatacaatccaaatcc	6600
Db	6541	ATAATCAACAAGATATGGGATCGCGCGCGCTGATCTCATTTCCATATCCAAATTC	6600
OY	6601	tgaagcttgactatgcccctacgatactgcgcacatccgfrgcccgtatattgaagcgtaac	6660
Db	6601	TGAAGCTGACTATGCCCTACGATCGCATGCCACCATCGGCTCGGTATTGATGAGCTGAAC	6660
OY	6661	cgacctaactgctgtcccgcaatccaaacgcacctatgtgtgtctgtcccgcaagctcgt	6720
Db	6661	CGCACCTACCTGCTGCCCGGAATCCAAACGCCCTATGTGTGTGTCGCGGACAGTGGT	6720
OY	6721	gcgcgtctgactctctgttgaaaggtatccgatcgaatcgaatctgattctgtggcggtg	6780
Db	6721	GCGCTGTTCACACTCTCTGTTGAAAGGCGCTATGCCATCGACGACGCTTCATGTTGGGCGGTG	6780
OY	6781	gaacgaagacccaagcgagagcgtgcgaatcctcaaatcatatgaatataatccgccgaa	6840
Db	6781	GACGAGGAGACCAAGGACAGCTGCCATGCTCAAAATTCATGAAGAAATATTCGGCCGAA	6840
OY	6841	atgatagaacgagccgacttaatctgcgaacggcgcgatgcgtctctgtctcgaacgaag	6900
Db	6841	ATGATGAACGCGCGACTAATGCGGACGGGCGGACAGTGTCTGTCTCGACGAGACAG	6900
OY	6901	cagcgatlgactcactatgtttgatalcgcgctgtgcgcgcacatctctgtgaaagccaaac	6960
Db	6901	CAGCGATGACCTCTACTGTTGATCGGCTTGCGGCGACATCTCTGAAGAAAGCCAAACAC	6960
OY	6961	gcgcgcaacgcaacgcgcgcgcggtgtgtcggaaaaaagctatctatgactgtcttcgg	7020
Db	6961	GCGGCAACACGACACGCGGCGGTGTGTGGGCAAAACCTTATCTTAGACTCTGCTTGG	7020
OY	7021	gaaccgacatccgaagaactccgaatgaagaacgaagcaatgagttcggatcaacttaagtt	7080
Db	7021	GACCGGATCCGAAGACTCCGATGAAGAACGAAGCAATGATTTGGATCACTTACGCT	7080
OY	7081	ctctcttccatgaaagcgccgcgaccagcagccacagcaccactgattcaagtgtaga	7140
Db	7081	CTCTCTTCCATGAAGCCCGCCGACACAGCACAGCACCACTGTTCAAGGTGGGA	7140
OY	7141	tgcgactatgatatatacctaagaacgcgacgctgtaggggttgatttgcaagagtgaagcc	7200
Db	7141	TGCGACTGATTAATCTTACAGACGACAGCTGAGAGGGTATTATTCGAAGGTGAAGCC	7200
OY	7201	gaattcatcatgttgatcatcgtattggaagaagggctgcggaagacacacctgcgcgg	7260
Db	7201	GAATTCATCATGTGATATGGGATTGGGAAGAGGGCTTGGCGGAAGACACCTTGCGCCG	7260
OY	7261	ctgcacgcgctcatctggaagaacgagggcatloaacttlltcaatgaaacgcgacgactcg	7320
Db	7261	TTGCACGCGCTCATCTGGAAGACGAGGGCATCACTTTTCACTGAACGCGACAGTCG	7320
OY	7321	gttgatcgggtgtgcgacactctctgtgcgggcgaatagacggggcgcaaaaactgtcgag	7380
Db	7321	GTTGATCGGGTGTGCGACATCTTCTGTGCGGGCCAATGAAGAGGGGGGACGAACCTGTGCAAG	7380

QY	7381	gcagaccctgctgatalgtcgtatgatcatcagctcaaaaaatggtccagcggatcggcccgcgagaa	7440
Db	7381	GCACACCTGGCTGATGTCGATGATGATCACGTCAAAATGATGTCACGAGGATTCGGCCCGGAGAA	7440
QY	7441	atcgcgagcctttgtcgagcaataaacaagttctggtggtcgcaaaaagatcaatgcgc	7500
Db	7441	ATCGCGCGCTTTGTGTCGAGCACATAAACAAAGCTCTCGGCGCCGCACAAACATCAGTCGC	7500
QY	7501	gacctggtccctgaagggcctctctgctcgtctcgatatagtatgctgctctataatgtaacg	7560
Db	7501	GATCTGCTGCTCGAAGGCGCTTCTCTGTCGTCTCGGATATTAGATGTCGTCTATTAAATGTCAGG	7560
QY	7561	aacttcaaacagcgagatcatcggcagaagatcogaagaacatggaatcgtatcatcaaacagca	7620
Db	7561	AACCTTACAAAGCGAGTCATCGCGCAGAGATCGAAGAACACTGGAGATGATCAACAGGCA	7620
QY	7621	ttcgagaacacggtctcgctcgtctgcaacaggcatgtgcatcatccggggataaactcggtctc	7680
Db	7621	TTTCAGAAACAGCTTCCCGCTCGCTGAACAGGCAATGCGATCACGGGGATTAACCTCGGCTCT	7680
QY	7681	ttgaagcgggtgctgctcctctggtctctatactataacaacgcggatcttcgatcttcgga	7740
Db	7681	TTGGAACCGGCTGCTGCTCTCTGTCCTATTATATCTACACACGCCGGATTTGGATTTCCGA	7740
QY	7741	ggaatcgacgagttcgagcgggtlcaaigtccagctccatgcaactctgtgttggtgaacagc	7800
Db	7741	GGATTCGAGCGAGTTGCGAGCGGGTCAATGCGAGCTCCATGCATCTGTGTTGGTGAACGC	7800
QY	7801	ctgctgtgtcagcgctctcgtttgtgcattcgatcatagacaatacacaacgcgcgcaatacg	7860
Db	7801	CTGCTGTGTCAACCGCTTTCGTTGGCAATTGGATCAACACATCACCGCGCAATACGC	7860
QY	7861	atccgcgatacaacccgtgtgtaaggccgcgcatcttcccagtaacgaagcctgttcgatctgcat	7920
Db	7861	ATCCGCGATCACTCCGCTGTGTAAGCCCGCGATTTCCCGATACGAACCTGTTGATGTCATG	7920
QY	7921	gcgaagaggggagcggctatctcaggttgaatgaagcgtacatccatcgaagaaatgtcctggaatg	7980
Db	7921	GCGAAGGGGGAGCGGCTATCTCAAGTGAAGTAAGACCGTACGAAGAAATGTCGTGAATG	7980
QY	7981	caatatggaagcccggaacctcgtttgcgctgtgcgtctatcgaaggcatcgactgtg	8040
Db	7981	CAATATGGCAAGCCCGGACCTTTCGTTGGCGTGTGCTCTCTATCAAGGGATGACATGG	8040
QY	8041	aacggaatcgaccttgcaatgtgcataatcatctcccacggagacgcctcagaaaaaatgtg	8100
Db	8041	AACGGATCGACCTGCGATGTGTGCATCATATCTTCCCAAGCGGACGCTCAGAAAAATGTG	8100
QY	8101	ctgcgcgggcgcaaatctgcgccgagacatcgcatcagaaaatcttgggcgcggttacaagt	8160
Db	8101	CTGCGCGGGCGCATGTGCCGAGCATCGGATTCAGGAANAATTGGGGCGGTTAAACGT	8160
QY	8161	ttggggcaaccttgcaactcttgcgcggagatagaaataatcagaaaaagtgtgcctgcatctc	8220
Db	8161	TTGGGCAACCTGCAACTTTTGGCCGGAGATGAAATATGAAAAAAGTGCCCTGCATTC	8220
QY	8221	agatcatggaatttaccggaacggcgctttgattcttactgaagacagatatcatctccggcgac	8280
Db	8221	AGGTCATGAGATTACCGGACGGCGCGTTGATTTCTAGACACACCATATGATCCGGCGAC	8280
QY	8281	cttgaacatgtgcgaatgtacatgcatactgcgccagttcgtgtcggaaacgggaaaaagtgtgac	8340
Db	8281	CTTGAACGTGCCGATGTACGTACGATCTGCCGATTCGTGCGGAAGGGGAAAAGTGTACTG	8340
QY	8341	cggcgcgctttgatgtgaattgtgtcggagacgagacgcaatgaatgaagttcgtctgttcaac	8400
Db	8341	CGGCGCGCTTTGATGTGAATTGTGTGTCGAGACGACGACGCGCATGATAGAGTCTGCTTGTAC	8400
QY	8401	gcgaagagcttgcgtgcatactctgtctcgacctgtgtgaaaaaagcgcgttcgaaacaacctg	8460
Db	8401	GCGAAGAGCTGCGTCATCTTGTCTGACCTGCTTGAATAAAAGCCCTGTCAACACCTGTG	8460
QY	8461	cggaagaccgaagcaagctgcgcgcccgcgtatgttltgacccgcgacga 8509	







Db 1681 ATGATCTGGGCGATGGTATAGATCCTGCAGAAATGACAGAGATTTCCTCCATT 1740  
Qy 4695 caaagaaatagacggaacacactaaacgcaagatatagtgaaggtatgaacaactg 4754  
Db 1741 CAAGAAATATACCGAAACACCTTAACAGCAAGTATATGTAAGCGGTATGAACACATG 1800  
Qy 4755 aaatgacctcgaagaaatcagccatlagagctgtaagcttcgagcttcaacgagatgctat 4814  
Db 1801 AAATGACCTCAGAAATACGCGCATTTAGAGCGTGAAGTTCGGACTTTACGCCATGCTTTAT 1860  
Qy 4815 ctgattgcagaggagatctgcgcaacaaatgvcgacatagccgagcgtcttcaatlcacat 4874  
Db 1861 CTGATGCGAGAGAGATGCGCAACAATGCGGACATGCGCGAGCGTTCATTAATTTAT 1920  
Qy 4875 caccgatagaagagaaacacgccccctcaaaaacaaagtgtggaagatattctat 4934  
Db 1921 CACCGATAGAGAGAAACCGCCCCCTCAAAAACAAAGATGGTGAAGATATTCGAT 1980  
Qy 4935 cctggcttcagagaccccttgctttaaactgaatcagcatcttcagcagctgataag 4994  
Db 1981 CCGGGCTTCAGAGACCTTGCTTTAAACCTGAATCAGCATCTTAGCGATGCTGATAG 2040  
Qy 4995 aagtaaatatagcacaatagagcgccatttccatcacaatagcgtcatalgtgat 5054  
Db 2041 AAGTAAATATAGCCACAAATAGAGCGGCATTTCCATTACATACAGCTCATCATGTAT 2100  
Qy 5055 caataatcagatgatattatcaatgtagaagaattatcatatgatacagagatcaca 5114  
Db 2101 CAATATCAAGTATGATATTCATCATGAGAGAAATTTACATGATACAGATCATCA 2160  
Qy 5115 cagcaattgtcttctatcttctaaatgctaaactaactacgctgcgacctaaagaagt 5174  
Db 2161 CAGCATTTGTTTTGTATTCTTAAGTGTACATACATACATGCTGCTTAAAGAAAGATT 2220  
Qy 5175 gtaactatgcagtalccacttgggtttcagagccgataatcaggtcttgcgaagcgg 5234  
Db 2221 GTACTATATGACGATATCTACCTTGGGTTTCAGACCGAATATTAAGCTTTGCGAAGCGG 2280  
Qy 5235 gttttttgcagagaagagaaacgcttcttgatgtgtaataactgtcagatctcatg 5294  
Db 2281 GTTTTTCGCCAGAGAGACACGTTTTTGTATGGGTATTAACGTCTACGATCTCATNG 2340  
Qy 5295 atgttaacatacacagcatcaccgagcaacacacacccctcttaatggaagatgataatg 5354  
Db 2341 ATGTGAACATACACAGCATCATCCGGGCAACACACCTTATGGAAGATGATTTATNG 2400  
Qy 5355 gcccaagaatctgcgcggaaaaaacgcatattgaagcagagccgcgcatcttagagcg 5414  
Db 2401 GCCCGAAGTACTGCGGAAAAAACCGCATTTGAGGCAACAGCCCGCACTTTAGAGCGCG 2460  
Qy 5415 aagcctataagcagatcacaacaaactagaaagccagattgaggaagaaatagggaagctg 5471  
Db 2461 AAGCTTATACGAGTACCAAAACCTAGAAAGCCAGATTGAGAAATATGGGAACGTG 2517

RESULT- 4  
AX281573 2112 bp DNA linear PAT 02-NOV-2001  
LOCUS AX281573  
DEFINITION Sequence 1 from Patent W00177347.  
ACCESSION AX281573  
VERSION AX281573.1 GI:1608825  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct.  
artificial sequence.  
REFERENCE  
1 (sites)  
AUTHORS D'Elia,J.  
TITLE Kelogulonigenium shuttle vectors  
JOURNAL Patent: WO 0177347-A 1 18-OCT-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
Source  
1..2112  
/organism="synthetic construct"

/db\_xref="taxon:32630"  
/note="replcon from pADW291"  
BASE COUNT 530 a 509 c 605 g 468 t  
ORIGIN  
Query Match 23.6%; Score 2007; DB 6; Length 2112;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 2955 ggcacatggctcgaatctcaatagaaattgtgtgaaggtgcagtcggtcttcgacaggggtg 3014  
Db 1 GGCATATGGGTGCAAAATTCATAGAAATTTGTGAGAGTGGCGTACAGCGCTGCACAGGGGTG 60  
Qy 3015 ctgcgcggagatctctgtgtctcaagtagggcgacaatgtgaaggggtgttagttgcccccg 3074  
Db 61 CTGCGCGGAGATCTCTGCTCAGAGAGGCGCAACAATGAGAGGTGTATGTTGCCCTCG 120  
Qy 3075 tatcgctctctgcgtgcgtcagatctgcgtccctgcgcgcgacatatgatatctccgataag 3134  
Db 121 TATCGCTCTCTCGCGCGCATTTGGGTCAATCCTGCCGACATATGATATTTCCGCTAGAG 180  
Qy 3135 gattactgatagtcttcgcgtcgtgcgtgtgcgggtgtgtgcgggtgtgtcgggtgtgc 3194  
Db 181 GATTAAGTATGATTTCTGCTGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCTTGTG 240  
Qy 3195 gggcgtgtccctctgtcccgctgtccctcaatttttcaacatcaaaaaatggcggaagc 3254  
Db 241 GGGCGTGCCTCTGTTGCCGCTGTCTCACTTTTTCACAAATGAATGGCGCAACG 300  
Qy 3255 cctctctgtctcatagttctctatagttcaatagcaaaatcacataatcatcaatagctt 3314  
Db 301 CCTTCTTGTCTATAGTTCTTTATAGTTCAATACGAAATTTACACATATTTATCAATAGCTT 360  
Qy 3315 atctgccttaaaaggagtaatttggccgcaaaaggagtaatttggccgcaaaaggagat 3374  
Db 361 ATTCCTTAAAGGAGATTAATGGCGCCCAAAAGGAGATAATGGCCGCAAAAGGAGAT 420  
Qy 3375 aattggccgcaaaaggagtaatttggccgcatatcgtgtgtttcatattgggaagatcc 3434  
Db 421 AATTGGCGCGCAAAAGGAGTAATGGCGCGATTCGTTGTTTACATGGGAGGAAATCC 480  
Qy 3435 ccttaatcattctcccccattgggaagaacaacaagttgcgcgacagccggccttcgac 3494  
Db 481 CCTTATATCATTTCTCCCAATGGGAAAGACACAAAGTGGCGGACCGGCGCTTCGAGC 540  
Qy 3495 cagacaacaaactgtgtcctccctgcgaggttgcgcgagaggggtctatatatgcgaatccgcgc 3554  
Db 541 CAGACAAAACATGTGCTCCTCCCTCCGAGTGGCGAGAGGGGTCTATATGCGCAATCCGCGCC 600  
Qy 3555 ggcctgcgagcgctcaagctcatgttaattagatagccactgvcgvcgvcgcatatgct 3614  
Db 601 CCGCTGCAGCGCCTCAAGCTCATGCAATTTAAATGATAGCCACTGCGGCGCGCATAGGCT 660  
Qy 3615 gatgatgtgcgcataagaaatgcggtcgtgcgcgacatctgcgcaatcgaagcgaatgaagaac 3674  
Db 661 GATGATGTGCGCGCATGAATATGGCGCTGCGGCGCATTCGCGCAATCGACGCAATGAAGAAC 720  
Qy 3675 catgacgctgaaagcctgaccccgctgttcgagaagctagccgctgtgcgtgtgttgaaccaat 3734  
Db 721 CATGACCGTGAAGAGCTGACCCGCGTTCGAGAGACTAGCGGCTGCGTGTGACCCAT 780  
Qy 3735 gatgacctgcgaagaatgattgtgataagtcgvcgvcgctgtgtcgaatgaagcgaataagac 3794  
Db 781 GATGACCTGTGCAAAAGATGATCTGTACAGTGGCGGCTTGTGTGATGAGCGCGCAATAGAC 840  
Qy 3795 taacgcaagaggaagcagcgcaactcctcagtgagcgttgagacctccggggtacatccgt 3854  
Db 841 TACCGCGACAGAGCAAGCGGCAACTCTAAGTACGTGAGCTTCCTCCGAGTACATTCGCT 900  
Qy 3855 cgtatggcgcgagtgatcgaaacacttgggcattcttcgacgcgtcaaaaggtlatccatctc 3914  
Db 901 CCTATGGCGCGGAGTGAACCACTGGGCCATTCCTGACCGTCAACGGGTATTCATCTTC 960



QY 3915 ggtgaatgaatcgtgctgctgtccagcagctctctagctcgtccaatctgacg 3974  
 Db 961 ggtgaatgaatcgtgctgctgtccagcagctctctagctcgtccaatctgacg 1020  
 QY 3975 atgagcgcaaaaccccttccgctcccgagcttgcgagcctctctgagctcccgagga 4034  
 Db 1021 atgagcgcaaaaccccttccgctcccgagcttgcgagcctctctgagctcccgagga 1080  
 QY 4035 aagatgctcgttgaagcagcagcttcaacagatcttctcacaacctgactgagtgatc 4094  
 Db 1081 aagatgctcgttgaagcagcagcttcaacagatcttctcacaacctgactgagtgatc 1140  
 QY 4095 aaccattatcgcgtctgacatgaagcgcaaacccgacccaagtctgacgtagcgagga 4154  
 Db 1141 aaccattatcgcgtctgacatgaagcgcaaacccgacccaagtctgacgtagcgagga 1200  
 QY 4155 agtgtgactaagcgtgaggaagtgaaagacgacccaacgctgcacagcgcgagctgagc 4214  
 Db 1201 agtgtgactaagcgtgaggaagtgaaagacgacccaacgctgcacagcgcgagctgagc 1260  
 QY 4215 ggttccaagtcggtcagagctgctcgcagagggcgagcggaagacatagccctcc 4274  
 Db 1261 ggttccaagtcggtcagagctgctcgcagagggcgagcggaagacatagccctcc 1320  
 QY 4275 ttcccaagaagcgagcgagctacactacagctccagcttgcgtgagctgagcagcctgct 4334  
 Db 1321 ttcccaagaagcgagcgagctacactacagctccagcttgcgtgagctgagcagcctgct 1380  
 QY 4335 ggcagcaacaagcaacaagcagctcagctcagactccgagcttctcgtcggagaga 4394  
 Db 1381 ggcagcaacaagcaacaagcagctcagctcagactccgagcttctcgtcggagaga 1440  
 QY 4395 ggcgtgcgtctgagcagctcgaacacatcgaaaaactgtttttagatcttcgcaaaagta 4454  
 Db 1441 ggcgtgcgtctgagcagctcgaacacatcgaaaaactgtttttagatcttcgcaaaagta 1500  
 QY 4455 ggggaagcttgcgttgcagcttctcagcagcaatagtgtaaatcttcgtaaacg 4514  
 Db 1501 ggggaagcttgcgttgcagcttctcagcagcaatagtgtaaatcttcgtaaacg 1560  
 QY 4515 atgtgcaatagcgttaagaactatgaatacacgctgagcagctgcgaagaagcaagc 4574  
 Db 1561 atgtgcaatagcgttaagaactatgaatacacgctgagcagctgcgaagaagcaagc 1620  
 QY 4575 gttggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4634  
 Db 1621 gttggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680  
 QY 4635 atgaatctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4694  
 Db 1681 atgaatctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740  
 QY 4695 caaagaataacacgcaaac 4754  
 Db 1741 caaagaataacacgcaaac 1800  
 QY 4755 aatgaactcagaataac 4814  
 Db 1801 aatgaactcagaataac 1860  
 QY 4815 ctgattgcaagggagagcagcagcaaaatggcgagacatggcgagcgtcttccaattcat 4874  
 Db 1861 ctgattgcaagggagagcagcagcaaaatggcgagacatggcgagcgtcttccaattcat 1920  
 QY 4875 caccgatag 4934  
 Db 1921 caccgatag 1980  
 QY 4935 cctgggctcagagagccttgcctttaaaccctgaa 4969  
 Db 1981 cctgggctcagagagccttgcctttaaaccctgaa 2015

RESULT 5  
 LOCUS AX281575 5859 bp DNA Linear PAT 02-NOV-2001  
 DEFINITION Sequence 3 from Patent WO0177347.  
 ACCESSION AX281575  
 VERSION AX281575.1 GI:16608827  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 . synthetic construct.  
 . artificial sequence.  
 .  
 REFERENCE  
 AUTHORS D'Elia, J.  
 TITLE 1 (sites)  
 JOURNAL Ketogulonigenium shuttle vectors  
 Patent: WO 0177347-A 3 18-OCT-2001;  
 Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
 FEATURES  
 source  
 1..5859  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="PADM291-4DS"  
 BASE COUNT 1444 a 1487 c 1566 g 1362 t  
 ORIGIN  
 Query Match 23.6%; Score 2007; DB 6; Length 5859;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 2955 ggcgaatgggtcgaatcgaatcgaatcttctgtagagctgagcagcctcgaaggggtc 3014  
 Db 402 ggcgaatgggtcgaatcgaatcgaatcttctgtagagctgagcagcctcgaaggggtc 461  
 QY 3015 ctgcgcgagatctcgtgctcagtagagggcgacaatgagaggtgttagttgccccctg 3074  
 Db 462 ctgcgcgagatctcgtgctcagtagagggcgacaatgagaggtgttagttgccccctg 521  
 QY 3075 tctgcgtctcgtgagcagcttgggtacatcctgcgcgcagacatgatatctccgttagag 3134  
 Db 522 tctgcgtctcgtgagcagcttgggtacatcctgcgcgcagacatgatatctccgttagag 581  
 QY 3135 gattactgaatgttctgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3194  
 Db 582 gattactgaatgttctgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 641  
 QY 3195 gggcctgtccctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 3254  
 Db 642 gggcctgtccctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 701  
 QY 3255 cctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 3314  
 Db 702 cctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 761  
 QY 3315 atctgcctaaaggaatgaatctggtcgcgcaaaaggaatgaatctggtcgcgcaaaaggaat 3374  
 Db 762 atctgcctaaaggaatgaatctggtcgcgcaaaaggaatgaatctggtcgcgcaaaaggaat 821  
 QY 3375 aattgggcgcgcaaaaggaatgaatctggtcgcgcaaaaggaatgaatctggtcgcgcaaaaggaat 3434  
 Db 822 aattgggcgcgcaaaaggaatgaatctggtcgcgcaaaaggaatgaatctggtcgcgcaaaaggaat 881  
 QY 3435 ccttaatcattctccccaatgggaagaacaacaagaatggcgagacagcgagcagcgagcagcagc 3494  
 Db 882 ccttaatcattctccccaatgggaagaacaacaagaatggcgagacagcgagcagcgagcagcagc 941  
 QY 3495 cagaacaanaatctgtcctcgtcgcgaggtggcgagaggggtctatctatgcgaatcgcgcc 3554  
 Db 942 cagaacaanaatctgtcctcgtcgcgaggtggcgagaggggtctatctatgcgaatcgcgcc 1001  
 QY 3555 cgcctgagcagcgtcgaagctatgcatctaatgtagcagcagcagcagcagcagcagcagcagc 3614  
 Db 1002 cgcctgagcagcgtcgaagctatgcatctaatgtagcagcagcagcagcagcagcagcagcagc 1061



[illegible]

Db	Accession	Version	Keywords	Source	Organism	Reference Authors	Title	Journal	Pubmed	Reference	Authors	Title	Journal	Pubmed	Comment
Db	2142	CNAAAGAAATFACACCGAAACACCTTAACACGCAAGTATATGCTAAGCGTATGTAACACATG	2201												
Qy	4755	aaatgacccacgaatacagcgcataatgagcgttgaagtcggacattacgcgatgcttatt	4814												
Db	2202	AAATGACCTCAGAAATCAGCGCATATGAGACGCGTGAAGTTTCGACCTTACGCGATGCTTAT	2261												
Qy	4815	ctgatgcagagagagatcgcgcacaaatgagcgcacatgagcgcgcatttcattcat	4874												
Db	2262	CTGATGCCAGGAGGATGCGCGCAAAATGCGCGCATGCGCGAGCTCTTTCACAAATTCAT	2321												
Qy	4875	caccgataagagagaaacgcgcgcctcaaaaaaagaagtgtgaggaagtattcgtat	4934												
Db	2322	CACCGATGAGAGAGAAACCGCCCCCTCAAAAACAAAGATGTGGAAGATATTCGTAT	2381												
Qy	4935	cctgagcttcagagacgttccttcaaaacgttaa	4969												
Db	2382	CCTGGGCTTCAGAGACCTTGCTTACGCGCGAA	2416												
RESULT 6	AE008913/C	23457 bp	DNA	linear	BCI 25-OCT-2001										
LOCUS	Salmonella typhimurium LT2, section 221 of 224 of the complete														
DEFINITION	genome: AE008913.1 GI:16423088														
ACCESSION	AE008913.1														
VERSION	AE008913.1														
KEYWORDS	Salmonella typhimurium LT2.														
SOURCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.														
ORGANISM	1 (bases 1 to 23457)														
REFERENCE	McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Portwoll, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Milvaney, E., Ryan, E., Sun, H., Flores, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R.K.														
AUTHORS	Complete genome sequence of Salmonella enterica serovar Typhimurium LT2														
JOURNAL	Nature 413 (6588), 852-856 (2001)														
PUBMED	11677609														
REFERENCE	2 (bases 1 to 23457)														
AUTHORS	The Salmonella typhimurium Genome Sequencing Project.														
TITLE	Direct Submission														
JOURNAL	Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA														
COMMENT	Supported by NIH grant 5U 01 AI43283														

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: GLIMMER: <http://www.tigr.org/softlab/glimmer/glimmer.html> and Genemark: <http://opal.biology.gatech.edu/genemark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes: <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc: <http://ecocyc.org/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and ReguindB: [http://kinich.cifn.unam.mx:8850/bb/reguindb\\_intro.frameset](http://kinich.cifn.unam.mx:8850/bb/reguindb_intro.frameset)

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by sequence from more than one ml3 subclone.

Location/Qualifiers

1..23457







AEVITSQGLTYONRTGMANDRLKAGISLSSRGKMCITPAGMSWVASHPOPMTELEF  
SHFACPNVKSLLADAVLSDPLEDEDLARSSPDRLRQALNEIRESAEELL  
ENLQVSPAFEVIVDLVLRITGGHRLQVGGGIDGIIISLDKLGEXYV

Query Match 1.9%; Score 165.6; DB 1; Length 23457;

Best Local Similarity 56.2%; Pred. No. 2e-26;

Matches 334; Conservative 0; Mismatches 254; Indels 6; Gaps 1;

Qy	359	ttgtctctctgctgcatgagcagcagcagcgttcacagggctcatatccagac	418
Db	10011	TTTTGTCGGCGCATCTGCTCCGCAAGAAVGGATTGACCCGCTCTGTTACCTTTT	9952
Qy	419	ccgcgcgccttgagcctacgcatcgcatcaagccgcggtgatgctgcttg	478
Db	9951	CCTGGCCGTTTCAGCCACTCATCATCGGCTCAATACGACCGAGGCTGAGTTT	9892
Qy	479	atcgcgcatagaaagcccccagctcgctgcgccattgagctcgatcagc	538
Db	9891	AACGGCGGAAAAAATTACACAGCATGCTTCAGACCATCCCATCCTCCGAAGCGGA	9832
Qy	539	acatcgttgatgcatctgctcgagcagcagcgtccacgcgcgcgaattcttcg	598
Db	9831	ATGTCGCTGTAATCAATCTCATCTGATTTTTCGCCAGCGCTTCAGTTCAGCTCA	9772
Qy	599	cgctccctcgtaagagcgcgagcgtatcgaaagtgatcaaccattgcatatctt	658
Db	9771	TGCTGAGACACTCAGCGGAGCGCTTGCCGCTTATGTTTAACCATCTCTATACGAT	9712
Qy	659	ccttcctcgagtgtagcgcgcgcgcgcgaatgatgcgcgtctcgcgcgcgcgc	718
Db	9711	CCTCTCTTTTCTGTCGCTGCTGCGGCGCTGATGATAGATTAATCTCAAAACACCTTC	9652
Qy	719	gcgcgcctcgatgctcgggtgagcagcagcagcgcgcgcgcgcgcgcgcgcgc	778
Db	9651	GAACGCAAGCGCTA-----TGAGCAACCGAGATGACCAATGCAATGAAACAGCGCAG	9598
Qy	779	gatttcgcagcgtatcccgccctcgatcctatccctgaaccctgacggtgcaacg	838
Db	9597	GGTCTGCCAAGCATTTCTTCATTTCTCATACGCTCCGACGTGACAGGTGTCCGGGATC	9538
Qy	839	ggcgaacacatgacgcatcctcgaaacgcgcatgctctcttccttcgcttcgcgc	898
Db	9537	GTCAAACACACAGTACCGCATCTCAAAAACGAGACCGGCTTCCTCCAGGTTGCTTAGC	9478
Qy	899	cttgcgagatccacagataaacgcattctcatgycgaattatactacacat	952
Db	9477	CTTGTTCGCATCCATTCACATCCATCGCATGTGTTTCCCTAATTGTACAT	9424
RESULT	7		
LOCUS	AL627284/c	88037 bp	DNA linear BCT 25-OCT-2001
DEFINITION	Salmonella enterica serovar Typhi (Salmonella typhi) strain C718,		
ACCESSION	AL627284	AL513382	complete chromosome; segment 20/20.
VERSION	AL627284.1	GI:16505618	
KEYWORDS	Salmonella enterica subsp. enterica serovar Typhi.		
SOURCE	Salmonella enterica subsp. enterica serovar Typhi.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.		
REFERENCE	1 (bases 1 to 88037)		
AUTHORS	Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,		
	Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,		
	Sedatalla,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T.,		
	Connerthon,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,		
	Farrar,J., Feltwell,T., Hamlin,N., Haque,A., Hien,T.T., Holroyd,S.,		
	Jagels,K., Krogh,A., Larsen,T.S., Leather,S., Moule,S., O'Garra,P.,		
	Perry,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J.,		
	Stevens,K., Whitehead,S. and Barrell,B.G.		
TITLE	Complete genome sequence of a multiple drug resistant Salmonella		
JOURNAL	enterica serovar Typhi C718		
	Nature 413 (6858), 848-852 (2001)		

PUBMED	11677608
REFERENCE	2 (bases 1 to 88037)
AUTHORS	Parkhill,J.
TITLE	Direct Submission
JOURNAL	Submitted (25-OCT-2001) Submitted on behalf of the Salmonella sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT	E-mail: parkhill@sanger.ac.uk Notes: Details of S. typhi sequencing at the Sanger Centre are available on the World Wide Web. (URL, <a href="http://www.sanger.ac.uk/Projects/S_typhi/">http://www.sanger.ac.uk/Projects/S_typhi/</a> ).
FEATURES	Location/Qualifiers
SOURCE	1..88037
	/organism="Salmonella enterica subsp. enterica serovar Typhi"
	/strain="C718"
gene	/db_xref="taxon:90370"
	482..1090
CDS	/gene="STY4858"
	482..1090
	/gene="STY4858"
	/note="Similar to Streptomyces coelicolor hypothetical protein SC7H1.19 TR:054178 (EMBL:AL021411) (182 aa) fasta scores: E(): 1.4e-10, 29.6% id in 169 aa"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="CAD03347.1"
	/db_xref="GI:16505619"
	/transl_table=1
	/note="N-terminus is similar to Salmonella typhimurium IICA IICA TR:09RP6 (EMBL:AF164435) (74 aa) fasta scores: E(): 3e-27, 98.6% id in 74 aa, indicated by promoterless reporter gene insertion to be induced upon bacterial invasion of phagocytic and nonphagocytic cells (but having no effect on virulence). The adjacent 70 aa is highly similar to the neighbouring CDS Salmonella typhimurium hypothetical protein TR:09RP5 (EMBL:AF164435) (70 aa) fasta scores: E(): 1.9e-20, 91.4% id in 70 aa"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="CAD03348.1"
	/db_xref="GI:16505620"
	/transl_table=1
	/note="N-terminus is similar to Salmonella typhimurium TLTYNHSNDSDEVYRRICKTILRKEYGELTYLAHSTYIDENSLNLTDDRKRIDITTY ELAVNIDTYLTLYTEVTLGDEPRDAISLNIFHEHSKTYTENGRVLRLKRLKLN ICYLVSGEGEYVDLMEAVNRLMLNDFACVVRCHAVRSVLRHHERKSTVHEQ ODSNTLTAMLMARGVISAEVKHFHFLIDVLESEIFLPAPERSHHNHAMLRISSEEE LACSIVYLTERRISGFLLQRAEDGAVP"
gene	2573..3034
CDS	/gene="STY4860"
	2573..3034
	/gene="STY4860"
	/note="Similar to Neisseria meningitidis putative periplasmic protein NMA0993 TR:CA884263 (EMBL:AL162754) (159 aa) fasta scores: E(): 1.2e-12, 34.4% id in 160 aa and Neisseria meningitidis hypothetical protein NNM0783 TR:AAFI1196 (EMBL:AE002432) (159 aa) fasta scores: E(): 1.5e-12, 34.4% id in 160 aa. Also similar to Haemophilus ducreyi hypothetical protein TR:AF63778 (EMBL:AF219260) (149 aa) fasta scores: E(): 9.2e-11, 31.5% id in 146 aa and Treponema pallidum hypothetical protein TP0895 TR:083865 (EMBL:AE001259) (168 aa) fasta scores: E(): 6.9e-09, 33.6% id in 134 aa. Contains a probable







QY 899 ctggcgagatccagataaacgcacatctcatg 932  
1111 111111 1111 1111 1111  
Db 29034 CTTGTGCAATCCATTCACAACTCCATCGCATG 29001

RESULT 8  
AE005851/c  
LOCUS AE005851  
DEFINITION Caulobacter crescentus section 177 of 359 of the complete genome.  
ACCESSION AE005851  
VERSION AE005851.1 GI:13423191  
KEYWORDS  
SOURCE  
ORGANISM  
Caulobacter crescentus.  
Caulobacter crescentus.  
Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
Caulobacter  
1 (bases 1 to 10661)  
Niernan, W.C., Feldilyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidberg, J.F., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwin, M.L., Haft, D.H., Kolonay, J.F., Smt, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L. and Fraser, C.M.  
Complete genome sequence of *Caulobacter crescentus*  
Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)  
21173698  
2 (bases 1 to 10661)  
Niernan, W.C., Feldilyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidberg, J.F., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B., Laub, M.T., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwin, M.L., Haft, D.H., Kolonay, J.F., Smt, J., Craven, M., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Shapiro, L., Venter, J.C. and Fraser, C.M.  
Direct Submission  
Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
FEATURES  
source  
location/qualifiers  
1..10661  
/organism="Caulobacter crescentus"  
/db\_xref="taxon:69394"  
complement(105..731)  
/gene="CC1766"  
complement(105..731)  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAK23742.1"  
/db\_xref="GI:13423192"  
/translation="MAQFDLPPKRPMTGRIIGPKPPKPHHIGIROOLKANGRV  
RDIAFNCAIDAKRGCDLKLKSDVAPGSLARATIIIOQGTGVPPEITPTPD  
ALAVWLKVRGRNDMLFSSRSRPGDHVSTRROYARLVDKMSMELEPRAYGHSLSLR  
TKVALLTKKSGNPAQCLLGHKRLSTVYRLGIVDALDKSEIDL"  
complement(1014..1547)  
/gene="CC1767"  
complement(1014..1547)  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="response regulator"  
/protein\_id="AAK23743.1"  
/db\_xref="GI:13423193"  
/translation="MAEPRRLIIEDEKPAATLKRSFERRGYEVANHGLQALETL  
RFTPRYAVVDLKGSGIACVRAALAEKRPDMILYVLTGPAATAVEAKICAGCHY  
LAKPSNTDIEAFLKREGDIDAIIGRATSIKLEMERIHQTLVETDFINSETARAL  
GMHRTTLARKLEKRRMS"  
complement(1540..2886)  
gene

CDS  
/gene="CC1768"  
complement(1540..2886)  
/gene="CC1768"  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="sensor histidine kinase"  
/protein\_id="AAK23744.1"  
/db\_xref="GI:13423194"  
/translation="MPPTANLVLSGSGMLSGAARAENANGPADATAROMLQILHR  
WIAVAGVATITLVHSESGFPPLAAMMLVLAIVNLISLRKRVBPGRELELA  
ALADALATITLYESGATNPFTTLLVLAIVLGAVALLEAATAVALISFEFVALI  
GFNRPIIAPGLSDRAPEFLYIMGMVLGLDVALVLFVGRINTNLSHARLAELO  
REAEFAHIRMGLASGAHELGPLATVLLIDVMDVRPAISNAELAELEDGCE  
VRRCIAITVILLASGARGSEVSTLRGMDEVAEMRASRADAPLIVETDLRDAT  
PIIAESTLKOGIHNILDMALAEASSAARLSAATPDGIRIVYEDDGCPFSALADGC  
KPTNSTGRAGSGIGLFLVNLKGLGVAEALVCGARVSTRPLVAMLEVARHG"  
complement(2883..3560)  
/gene="CC1769"  
complement(2883..3560)  
/note="identified by match to PFAM protein family HMM  
PF02104"  
/codon\_start=1  
/transl\_table=11  
/product="Surf1 family protein"  
/protein\_id="AAK23745.1"  
/db\_xref="GI:13423195"  
/translation="MLFAGLGVLWOLQRRVWKLDLIAOVBORLAAPRYGACPDLPPL  
APANDVRRVLSGVFDHRTTLQAVTVLGGPGWVLTPLRTDGFVLVNRGVPE  
RAASRRAGQVREIRVGLLRFTEGPEGLLRNQGPAAGSRVDAIAOSGGLG  
VAPFVADADAPNCGWPRGLIVVRRPNSHLIYALWFMALAEFSAGAAVVLADARR  
SRREDAA"  
complement(3617..4006)  
/gene="CC1770"  
complement(3617..4006)  
/note="similar to GB:X78196 PID:461330; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="ubiquinol oxidase subunit IV"  
/protein\_id="AAK23746.1"  
/db\_xref="GI:13423196"  
/translation="MSAASHDAHPDHGDAHEAAGTKLDQVYIGFVLAIVTPRPW  
LYMGDVPKSNQMTAVAYAGLAVIOVVYHMTIFLMMTRSEGGTMALITFLIIVIT  
LAGSLWVHNLNTHMPKEMPMKROAP"  
complement(4003..4629)  
/gene="CC1771"  
complement(4003..4629)  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="ubiquinol oxidase subunit III"  
/protein\_id="AAK23747.1"  
/db\_xref="GI:13423197"  
/translation="MGAHSQAVSAADPDQYLTREHNPENKTLIGFVIYLSQCLIRA  
VLRACVAVLOSIVAGPSGADLPDLKLVATIKLSSITTYGFAAMRAQAKREVPLV  
IWLGVTLGVLGFLSLIYEFANLIRHGATPORSAPLSAFLLVGTGLHVTGCVIWL  
VTLMVQVARGGLGIEMRRLMCLSMFWHFLDVIVIGVSFVYLLGVLK"  
complement(4632..6629)  
/gene="CC1772"  
complement(4632..6629)  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="ubiquinol oxidase subunit I"  
/protein\_id="AAK23748.1"  
/db\_xref="GI:13423198"  
/translation="MSPDLVQVIFGRLTWDAIPLHDPILLVTPFAVVALGVTVALLT  
RYKLIMGVLMREWLTSVDHKIKIGVYIILALIMLRGFADALMRAQOAMAFGATGYL  
complement(1540..2886)  
gene







gene /db\_xref="taxon:382"  
/plasmid="pSymba"  
complement(85..561)  
/gene="Sma1867"  
complement(85..561)  
/gene="Sma1867"  
/function="Miscellaneous; Not classified regulator"  
/note="glimmer prediction. Contains a motif similar to  
AsnC family. Similar to LRP\_KLEAE LEUCINE-RESPONSIVE  
REGULATOR Protein"  
/codon\_start=1  
/transl\_table=11  
/product="Putative AsnC (Lrp)-family transcriptional  
activator"  
/protein\_id="AAK65685.1"  
/db\_xref="GI:14524176"  
/translation="MKLDRIKILYELQKNGITIVELAEVLNLSPSPECLMRVKLQ  
SEGYIDGSAQINVKLGQTLVETITLKNRQIDFARFLAIEKVDVTECHLVSG  
GYDYLKFTVAGINEXQTIMERLTDMDVGIDKXFSFVLKSPVAKHMLTSLFRV"  
797..2122  
/gene="Sma1869"  
797..2122  
/gene="Sma1869"  
/function="Small molecule metabolism; Degradation; amino  
acids"  
/note="glimmer prediction. Similar to agae [Agrobacterium  
tumefaciens]. ACCESSION AAB88469"  
/codon\_start=1  
/transl\_table=11  
/product="Putative deaminase"  
/protein\_id="AAK65686.1"  
/db\_xref="GI:14524177"  
/translation="MPAPLKFVHTTCELPDAADVIVGGIGIFSAYYLARGLTKVA  
LVEKIGIAQOSRMKRCQONRDAELPMKSLDLMRFAKNGCPTGPRCGLF  
YLSNSEEELAGMARRODFRTVYTHMLDSADTERGRATGPKMGVSPSTGDTAD  
PASAPAVARAILLEGGTTHQSCARGITEGGSLGVVTEHGAIRTKAILAGGMA  
SSFCROLGRFQALRSSILAVSPGVTSLPDALHTAAVSTRSDGYSLAISGRRA  
IDPTPOOLRFAPOLPMFVKRMRSIAPGLGEGRSCHETLVRRLLDAPPMEMRIL  
PAVDNATIRLTHSRALDILPAKNTRTITAAAGYIDSTPDGVPGIGEIAAIPFIILAA  
GFSHGFGIGPCGAGLADIYVCSPIVDPHPHPRFSGKMGKVADE"  
2321..3349  
/gene="Sma1871"  
2321..3349  
/gene="Sma1871"  
/function="Small molecule metabolism; Degradation; amino  
acids"  
/note="glimmer prediction. Similar to C69066 339 aa  
ornithine cyclodeaminase - Methanobacterium  
thermautotrophicum (strain Delta H). ACCESSION C69066"  
/codon\_start=1  
/transl\_table=11  
/product="Putative deaminase"  
/protein\_id="AAK65687.1"  
/db\_xref="GI:14524178"  
/translation="MHPMPKTIERSCKMTLLKKDEVRLIGMAEYIGAVEAYKA  
SVOEQPDYIGIHHPISLGEIDFKLGYKANEIISMAKASGFTNPAEYVNSIGT  
ILLEFARSCALICINDGSLITGLRTGASGVSKALARKNARFASIGGNQMRQIR  
AVNEIMKIKIHAWSRPSISRYKTGDIORERCIIVASAKREAEQDILTTTGR  
KSLVADAWKRGSTHVAIGTDGRQOEIDPEIFRNKAYVDSVAACTEGEFTWHPLN  
KNITIKDDIHGEIGEVILGRKPGRESDETTITPDSMAIQDNTASKIYONAIANNV  
GTFQFFE"  
3355..4326  
/gene="Sma1872"  
3355..4326  
/gene="Sma1872"  
/function="Small molecule metabolism; Degradation; amino  
acids"  
/note="glimmer prediction. Contains motif similar to  
pyridoxal-phosphate dependent enzymes. Similar to D72386  
401 aa threonine dehydratase catabolic - Thermotoga  
maritima (strain MSB8). ACCESSION D72386"  
/codon\_start=1  
/transl\_table=11

gene /product="Putative threonine dehydratase"  
/protein\_id="AAK65688.1"  
/db\_xref="GI:14524179"  
/translation="MRRYPTIIDIIRERARELKHVRHTPLBRAKTEKAAACGLYLP  
ETLQITGAFKIRGALNKAULSREELANGIITAVSSGHHAGLSYAAAMLVKYLIVP  
VTPPKIKIENTKALGAELILEFDNNAARMKVYELAEKNVAVIHGEPEDVNAAGGT  
ICEEILDELDVDVTVPLGGGILSIGIAIETKRSVIAEPAITPKYFSRVN  
KERTSLPKWTIDGLRISVGGNPNPITIEKYDEIVLVEDEHIGMRALANDAKLI  
APPAASIGVALLAGIIDVKLDEKVCALVSGMWDRLDAIYIVAG"  
4616..4957  
/gene="Sma1874"  
4616..4957  
/gene="Sma1874"  
/function="Miscellaneous; Not classified regulator"  
/note="glimmer prediction. Local similarity to  
Integrase/recombinases."  
/codon\_start=1  
/transl\_table=11  
/product="Hypothetical protein"  
/protein\_id="AAK65689.1"  
/db\_xref="GI:14524180"  
/translation="MTDQTEAVGAMTIESRRLGEROYLPFSRYHTKPHLSTROYSRV  
ERNVSSIGLDPKRYGTHSMRRTVAHITKRTGNLRAVOLLGHHKLESTVQYLTGYD  
DALAISEQVEL"  
complement(5568..7085)  
/gene="Sma1875"  
complement(5568..7085)  
/gene="Sma1875"  
/function="Elements of external origin; Transposon-related  
functions"  
/note="glimmer prediction. Contains a motif similar to  
Reverse transcriptase (RNA-dependent DNA  
polymerase) similar to AAB6996 431 aa reverse  
transcriptase-like protein [Shigella flexnerii]. ACCESSION  
AAB6996"  
/codon\_start=1  
/transl\_table=11  
/product="Putative reverse transcriptase"  
/protein\_id="AAK65690.1"  
/db\_xref="GI:14524181"  
/translation="MDTDHRTDVMVWLGIGIQRKLYQMSKANPDQMDQWMLTDLRYL  
RHAMORVASKNGGRTAGVDMTVGRIENRSEHFLVDLQADLRSGAVRPSPARKLLP  
KAGKPCQFRPLGIPTIRDVYQGAALKLLEPIEADQFVGFRRPGRNHALLEYIR  
RALPQKRODEDTRRNRLRPVWVIEGDKGCEFDNNHHLERKRKRGDRVRLVGL  
FLKAGVTEPOPLRTDAGTPGGCISPLANINLSAIEEYEPVYTKRKTQARRSN  
GVAAASARDSDIACGYLTPRYADDPVYVLSIEEPAEKSAALADLITITGIL  
LPEKTKVTAMTEGFEFLGFRFSVHMDKRYGPRVEIPAKANLRHKVKQLTORDS  
ISVLSGEKRGVNAITSGMANVYRCVAGARVGVADLPMYIGRLYCMILHKRKPAPS  
ELWGSKOPSRRRATRRVREGSVHVGWTVPDVRYLAWMDPDMFAMSSGEPDA"  
complement(7697..8324)  
/standard\_name="ISKm25 fragment"  
/note="Sma4012; predicted by homology; local similarity to  
IS1407 of Mycobacterium celatum. 928 identical to ISKm25  
of S. meliloti pSymba. Probably a nonfunctional IS element"  
/rpt\_family="ISKm25"  
/evidence="not experimental"  
complement(7750..7905)  
/gene="Sma1878"  
complement(7750..7905)  
/gene="Sma1878"  
/function="Elements of external origin; Transposon-related  
functions"  
/note="glimmer prediction. Local similarity to  
transposases; likely non-functional"  
/codon\_start=1  
/transl\_table=11  
/product="Putative transposase"  
/protein\_id="AAK65691.1"  
/db\_xref="GI:14524182"  
/translation="MAENALTAVIQEAVYIGISTRSVDLVKGGIEMHLQDEGQRAV  
KGRDRQG"  
8432..9025  
/gene="Sma1881"















AUTHORS Freilberg,C.  
TITLE Direct Submission  
JOURNAL Submitted (22-NOV-1996) Genome Analysis, Institute for Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
REFERENCE 3 (bases 1 to 9973)  
AUTHORS Freilberg,C.  
TITLE Direct Submission  
JOURNAL Submitted (12-DEC-1997) Genome Analysis, Institute for Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
REMARK update by submitter  
FEATURES  
source  
1. .9973  
/organism="Rhizobium sp. NGR234"  
/plasmid="pNGR234a"  
/strain="NGR234"  
/db\_xref="taxon:394"  
/clone="pXB1423"  
<1. .3474  
/organism="Rhizobium sp. NGR234"  
/db\_xref="taxon:394"  
/clone="pXB686"  
62. .79  
/function="conjugal transfer terminator"  
/rpt\_type="inverted"  
/evidence="not\_experimental"  
81. .92  
/note="region including origin of transfer (oriT) almost  
identical to oriT regions of plasmids from the 'Q-group'"  
/note="Region: origin of conjugal transfer"  
/evidence="not\_experimental"  
238. .546  
/gene="trac"  
238. .546  
/gene="trac"  
238. .546  
/note="probable conjugal transfer protein"  
/codon\_start=1  
/transl\_table=11  
/product="Trac"  
/protein\_id="AAB92440.1"  
/db\_xref="GI:2182364"  
/translation="MKKPAKIRDEIKKLOBOLQATREAREIRGRIALAKGLGIEY  
EEAEIQAIFESLAKRFKGGGANGASSGGKKPDGNATGASAAVSGAAGGNGEA  
"  
551. .766  
/gene="trac"  
551. .766  
/gene="trac"  
/note="probable conjugal transfer protein"  
/codon\_start=1  
/transl\_table=11  
/product="Trac"  
/protein\_id="AAB92441.1"  
/db\_xref="GI:2182365"  
/translation="MARSATSDARKKDTREKIELGLIVKAGLREKRVLLGALVDL  
SRRUNDESERARLIAIGAEAFGDDG"  
753. .2675  
/gene="trac"  
753. .2675  
/gene="trac"  
/note="probable conjugal transfer protein"  
/codon\_start=1  
/transl\_table=11  
/product="Trac"  
/protein\_id="AAB92442.1"  
/db\_xref="GI:2182366"  
/translation="MTVNRIALVAVPVLMVLVIGNSGIEQCLSTFGKTDARLALG  
RIGAPPYVTAAAVLAVLFASAGSALITKLGWALAGSAAATMLIATMREASRLAA  
AAYAGKSIASYLDHATLIGATALMSGCSLRYALLIGNAFAFAEPPRIKRGKALG  
EADMNMNODARLESDAGVIGERIVVDKSTVARSFRAADDPETWAGGKSPGLCD  
GSPGSSHGIVFAGSGGKRTSVTPTALKKGALVIVDPSENAVMYERHKAARTI  
RIIDPKPETGFNALDRIIRGGTKEEDIASVAVISDSGMRGVDDFFRASALDL  
LTALIAVCLSGHTDKENQTLROYRANLAEPEPLRLRLQSIYDNSNSDFVKEVNAVY  
VNMPTETSGVYANAVKETHMLSYPNYAAVSGSTFSDALAAGETDVFINIDIKTLE

gene  
/note="fragments homologous to Agrobacterium tumefaciens  
ORFL1 (conjugal transfer region 1)"  
/pseudo  
3224. .3838  
/gene="y4dw"  
3224. .3838  
/gene="y4dw"  
/note="hypothetical 22.9 kd protein"  
/codon\_start=1  
/transl\_table=11  
/product="y4dw"  
/protein\_id="AAB92443.1"  
/db\_xref="GI:2182367"  
/translation="MRALPGAGTVCODTAVTERSERENRHPTRGPRAGSVASRAVR  
HRLAQAHSINISDMNRSVPTDSQSLDAIVGTSDDSLMTSOPRYLYVEHDSRNM  
ARYAMSTIEPMLFEDICILRKWGRIGTGQMMVHHFGQEDAVRLFDLILQKKRGT  
RPRSLPEKMPAEAEHSATSAVPDPBGHLDLDEERNLDGL"  
3863. .4441  
/gene="y4dx"  
3863. .4441  
/gene="y4dx"  
/note="hypothetical 20.6 kd protein"  
/codon\_start=1  
/transl\_table=11  
/product="y4dx"  
/protein\_id="AAB92444.1"  
/db\_xref="GI:2182368"  
/translation="MAAHLICRGAVGRREMLRSGKKPIRIAKAPAAATASSTSEY  
RATTATGELPAGDVRPHSPRTESQVLSDDADJGERPAARDELAEBGKEFPAEPD  
RLGLIDVKGANDIGVERIEDDPAQVCIAGPPLLDGRNFSRAASILGLLGDSREKQ  
NOVVIRSGKPHADGDWPLPPLVLSFSLPRA"  
4489. .4755  
/gene="y4ea"  
4489. .4755  
/gene="y4ea"  
/note="hypothetical 9.9 kd protein"  
/codon\_start=1  
/transl\_table=11  
/product="y4ea"  
/protein\_id="AAB92445.1"  
/db\_xref="GI:2182369"  
/translation="MPRPGSYSTSDLSRKSGDIIAEALRHPVITQRNKPRLVLLNI  
DDYERLMOYDARSVGTLETLPGLNFEFAAVYAVGTDENR"  
5209. .5209  
/gene="y4eb"  
5209. .5209  
/gene="y4eb"  
/note="hypothetical 11.6 kd protein"  
/codon\_start=1  
/transl\_table=11  
/product="y4eb"  
/protein\_id="AAB92446.1"  
/db\_xref="GI:2182370"  
/translation="MTMATIGTFTSTENGFTGSIITLALNVKARIGRIENPSDKGPH  
RIVAGAVELGAAMQKRSQDRDLYSVKLDDPSFAPIVATLLEVGEDGYOLISRPN  
RD"  
5748. .7926  
/gene="fei"  
/note="Fragments homologous to St. lividans mercuric  
reductase"  
/pseudo  
/evidence="not\_experimental"  
8224. .9147  
/gene="y4ec"  
8224. .9147  
/gene="y4ec"  
/note="hypothetical 34.2 kd protein"



```
/codon_start=1
/transl_table=1
/product="y4ec"
/protein_id="AAB92447.1"
/db_xref="gi:2182371"
/translation="MTDRTPTQTPDIERYTSOIIAIEAGASYDMPWHDSAIT
TVNVAASKAYRGVNIISLMAAAHAGIAGTGMCAIDAOVRKGERHLVFM
KTITDADADAGHDDREHPRRLFARGVTYFNCAOVGYTPPEMPVLPPEARIERAE
RCCASLGIDIRHGSQATYRSTDIYQHPERACPDAAVAYAVLLHEGSHGAKHRL
DRDLGRFESSAAYAMECTVELLSMTICADULSLVEARPDARIASLSLELRSDQRA
IFTAASKAQIADMMHAAQSDAQRNEVGGAA"
BASE COUNT      1844 a      3001 c      3352 g      1776 t
ORIGIN

Query Match      0.8%; Score 65; DB 1; Length 9973;
Best Local Similarity 55.6%; Pred. No. 0.0013;
Matches 125; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

OY 1809 aaccatgagacctatctcgagaagcgcatcccccgaagaacacagcaggttctacc 1868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3483 AACCTACCGGCTCTATGTCGAGCGCCTCGATCCCTCGAGGAATATGCGCCGCTACTATG 3542
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1869 gcatgcgagtcctcgcaacctgttcggggaatggaactgtatcgagaatgggcccga 1928
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3543 CCATGTGATCGAACCGAACCTCTTCGGCGACATCTGCTGCTTCGCAAGTGGGGTGGCA 3602
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1929 ttggcatcagcgccgcgcctcgcgtcgatgtgttgagagtgaaagaatgccaatcgctg 1988
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3603 TTGGAAACCAAGGACAGATGATGTCATCTTCGCGCAGAGAGAGATGCGGTCCGAT 3662
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1989 cgatgctgcgcacgcagacgcgcagcgatcagcgcggtatggc 2033
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3663 TGTTCCTGATCTGCTCGACAGAAACGAAACGCGGTTATCGTC 3707
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
LOCUS      A79350      320040 bp      DNA      linear      PAT 20-OCT-1999
DEFINITION      Sequence 1 from Patent EP0818465.
ACCESSION      A79350
VERSION      A79350.1 GI:6092393
KEYWORDS
SOURCE      unidentified.
ORGANISM      unidentified.
REFERENCE      1 (bases 1 to 320040)
AUTHORS      Broughton,W.J. and Perret,X.P.
GENOMIC SEQUENCE OF RHIZOBIUM SP. NGR234 SYMBIOTIC PLASMID
PATENT: EP 0818465-A 1 14-JAN-1998;
JOURNAL      BIOLOG MOLECULAIRE DES PLANTES (CH); INST OF MOLECULAR BIOTECHNOLOGY
      (DE)
FEATURES
      source      1..320040
                  /db_xref="taxon:32644"
BASE COUNT      66253 a      93201 c      94558 g      66028 t
ORIGIN

Query Match      0.8%; Score 65; DB 6; Length 320040;
Best Local Similarity 55.6%; Pred. No. 0.0019;
Matches 125; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

OY 1809 aaccatgagacctatctcgagaagcgcatcccccgaagaacacagcaggttctacc 1868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83564 AACCTACCGGCTCTATGTCGAGCGCCTCGATCCCTCGAGGAATATGCGCCGCTACTATG 83623
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1869 gcatgcgagtcctcgcaacctgttcggggaatggaactgtatcgagaatgggcccga 1928
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83624 CCATGTGATCGAACCGAACCTCTTCGGCGACATCTGCTGCTTCGCAAGTGGGGTGGCA 83683
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1929 ttggcatcagcgccgcgcctcgcgtcgatgtgttgagagtgaaagaatgccaatcgctg 1988
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83684 TTGGAAACCAAGGACAGATGATGTCATCTTCGCGCAGAGAGATGCGGTCCGAT 83743
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1989 cgatgctgcgcacgcagacgcgcagcgatcagcgcggtatggc 2033
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83744 TGTTCCTGATCTGCTCGACAGAAACGAAACGCGGTTATCGTC 83788
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
LOCUS      AF429315/c      125020 bp      DNA      linear      PRI 18-JAN-2002
DEFINITION      Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION      AF429315
VERSION      AF429315.1 GI:17646244
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 125020)
AUTHORS      Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL      Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE      21583737
PUBMED      11694876
REFERENCE      2 (bases 1 to 125020)
```

```
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83684 TTGGAAACCAAGGACAGATGATGTCATCTTCGCGCAGAGAGATGCGGTCCGAT 83743
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1989 cgatgctgcgcacgcagacgcgcagcgatcagcgcggtatggc 2033
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83744 TGTTCCTGATCTGCTCGACAGAAACGAAACGCGGTTATCGTC 83788
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
LOCUS      A93002      320040 bp      DNA      linear      PAT 22-JAN-2000
DEFINITION      Sequence 1 from Patent WO98023560.
ACCESSION      A93002
VERSION      A93002.1 GI:6741430
KEYWORDS
SOURCE      unidentified.
ORGANISM      unidentified.
REFERENCE      1 (bases 1 to 320040)
AUTHORS      Broughton,W.J. and Perret,X.P.
GENOMIC SEQUENCE OF RHIZOBIUM SP. NGR 234 SYMBIOTIC PLASMID
PATENT: WO 9802560-A 1 22-JAN-1998;
JOURNAL      BIOLOG MOLECULAIRE DES PLANTES (CH); BROUGHTON WILLIAM JOHN (CH)
      (CH)
FEATURES
      source      1..320040
                  /db_xref="taxon:32644"
BASE COUNT      66253 a      93201 c      94558 g      66028 t
ORIGIN

Query Match      0.8%; Score 65; DB 6; Length 320040;
Best Local Similarity 55.6%; Pred. No. 0.0019;
Matches 125; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

OY 1809 aaccatgagacctatctcgagaagcgcatcccccgaagaacacagcaggttctacc 1868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83564 AACCTACCGGCTCTATGTCGAGCGCCTCGATCCCTCGAGGAATATGCGCCGCTACTATG 83623
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1869 gcatgcgagtcctcgcaacctgttcggggaatggaactgtatcgagaatgggcccga 1928
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83624 CCATGTGATCGAACCGAACCTCTTCGGCGACATCTGCTGCTTCGCAAGTGGGGTGGCA 83683
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1929 ttggcatcagcgccgcgcctcgcgtcgatgtgttgagagtgaaagaatgccaatcgctg 1988
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83684 TTGGAAACCAAGGACAGATGATGTCATCTTCGCGCAGAGAGATGCGGTCCGAT 83743
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1989 cgatgctgcgcacgcagacgcgcagcgatcagcgcggtatggc 2033
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 83744 TGTTCCTGATCTGCTCGACAGAAACGAAACGCGGTTATCGTC 83788
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
LOCUS      AF429315      125020 bp      DNA      linear      PRI 18-JAN-2002
DEFINITION      Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION      AF429315
VERSION      AF429315.1 GI:17646244
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 125020)
AUTHORS      Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
JOURNAL      Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE      21583737
PUBMED      11694876
REFERENCE      2 (bases 1 to 125020)
```







**THIS PAGE BLANK (USPTO)**

---







XX Claim 1; Fig 1A-E; 14pp; English.

PS The present sequence represents a purified nucleic acid of a  
CC Ketogulonipentium plasmid endogenous to strain NRRL B-30035. The  
CC polynucleotide can be used in a vector. The vector comprises either  
CC (i) the polynucleotide and a marker gene or (ii) the polynucleotide, a  
CC promoter, a transcription terminator and a discrete series of restriction  
CC endonuclease recognition sites between the promoter and transcription  
CC terminator.

XX Sequence 8509 BP; 2065 A; 2239 C; 2345 G; 1860 T; 0 other:

Query Match 100.0%; Score 8509; DB 22; Length 8509;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 8509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggatgaacgcatctggttgatgtttgagaaagcgaagaacccggccacagttgtgggt 60  
Db 1 tggatgaacgcatctggttgatgtttgagaaagcgaagaacccggccacagttgtgggt 60  
QY 61 aagagcgcgatgtgtgaaagacactgacatgtgtgtgacatcgaaactcgtgtctacctgc 120  
Db 61 aagagcgcgatgtgtgaaagacactgacatgtgtgtgacatcgaaactcgtgtctacctgc 120  
QY 121 atcgtcgtatcatagcctgtctcgagatggcggaagaccgacctatgtgtcgtacagc 180  
Db 121 atcgtcgtatcatagcctgtctcgagatggcggaagaccgacctatgtgtcgtacagc 180  
QY 181 ccttaagcgatgugcggaacttggcccatgcccacatgctgtgttccacatcgccgagat 240  
Db 181 ccttaagcgatgugcggaacttggcccatgcccacatgctgtgttccacatcgccgagat 240  
QY 241 taacgaactggaagatgatactctagaagcggttaacccagacatctgtgttaacgcaataaag 300  
Db 241 taacgaactggaagatgatactctagaagcggttaacccagacatctgtgttaacgcaataaag 300  
QY 301 ctgcgcgatgaataagcgaacatactgtgcgtggccgcccgttctgaacgctgtct 360  
Db 301 ctgcgcgatgaataagcgaacatactgtgcgtggccgcccgttctgaacgctgtct 360  
QY 361 gctcctcctgcgtatgagacacgacagatcgctcatlacggtgtctgtatccagacc 420  
Db 361 gctcctcctgcgtatgagacacgacagatcgctcatlacggtgtctgtatccagacc 420  
QY 421 gccgcgcttgagcgaatgcagacacatcgcatcaagccgcggtgtatctgtctgtat 480  
Db 421 gccgcgcttgagcgaatgcagacacatcgcatcaagccgcggtgtatctgtctgtat 480  
QY 481 cgggcgatagaagcgccacagctcggtctgcccattgggcttcgcgtcagctcggaac 540  
Db 481 cgggcgatagaagcgccacagctcggtctgcccattgggcttcgcgtcagctcggaac 540  
QY 541 atcgttggatgacatctgctcggcgcgagagcgttccagccgcgccaattctctgtggcg 600  
Db 541 atcgttggatgacatctgctcggcgcgagagcgttccagccgcgccaattctctgtggcg 600  
QY 601 ctccctcggttaagagcgcgacgctatcgaaagttatccaacattggcatatctctcc 660  
Db 601 ctccctcggttaagagcgcgacgctatcgaaagttatccaacattggcatatctctcc 660  
QY 661 ttctcctcggttgaagcgcgacgcaaatatgtcgatcgtcttcgacccgagatcgggcg 720  
Db 661 ttctcctcggttgaagcgcgacgcaaatatgtcgatcgtcttcgacccgagatcgggcg 720  
QY 721 cagcctcgatgacaggtgggcaacacgaagagcgagcgccatagatctgtcccaacgg 780  
Db 721 cagcctcgatgacaggtgggcaacacgaagagcgagcgccatagatctgtcccaacgg 780  
QY 781 ttgtcagcgcgtatctccgcgacctgatactctgaacgctcaggtgtgcaacggtatcg 840  
Db 781 ttgtcagcgcgtatctccgcgacctgatactctgaacgctcaggtgtgcaacggtatcg 840

QY 841 cgaacacatgacagcatctctcgaaaccgagatgcattcttctgtctgtccgct 900  
Db 841 cgaacacatgacagcatctctcgaaaccgagatgcattcttctgtctgtccgct 900  
QY 901 tggcggatctccagataaacccgcatctcatgagagaattataactacaattgtatgt 960  
Db 901 tggcggatctccagataaacccgcatctctcatgagagaattataactacaattgtatgt 960  
QY 961 attcaatgcaagtcgcaggtttcaaatcacgcccacaacccgcaactgattctgttac 1020  
Db 961 attcaatgcaagtcgcaggtttcaaatcacgcccacaacccgcaactgattctgttac 1020  
QY 1021 tcaagcgcgcttttgaatagaagcttgatataaacacccgcgcttctcaacaaata 1080  
Db 1021 tcaagcgcgcttttgaatagaagcttgatataaacacccgcgcttctcaacaaata 1080  
QY 1081 aagcaaatccgcgcgcgtgagcaaatctgcgcttctgtatgcaaggtctgtgtttca 1140  
Db 1081 aagcaaatccgcgcgcgtgagcaaatctgcgcttctgtatgcaaggtctgtgtttca 1140  
QY 1141 tactgaaagacatgcaaggaattgcccgcgtgagcacacgaagacacccaagc 1200  
Db 1141 tactgaaagacatgcaaggaattgcccgcgtgagcacacgaagacacccaagc 1200  
QY 1201 ggccctgaaacgaagggcgcgctgttcgggaaagccgcgcgtgacacctgacagattgc 1260  
Db 1201 ggccctgaaacgaagggcgcgctgttcgggaaagccgcgcgtgacacctgacagattgc 1260  
QY 1261 cctgtatcgtctcatctcctgcgcgaagaaagcggtgggagatctgtgtcttcaagc 1320  
Db 1261 cctgtatcgtctcatctcctgcgcgaagaaagcggtgggagatctgtgtcttcaagc 1320  
QY 1321 gggatcgcgaacacagatttgcgcgcgtcgcgaacctcgtgcgcgtcgtcgtatgtgc 1380  
Db 1321 gggatcgcgaacacagatttgcgcgcgtcgcgaacctcgtgcgcgtcgtcgtatgtgc 1380  
QY 1381 gacccagctgtgtctgtgtgagatctcgagatccgcgaagaaagacccgaagcccgcaa 1440  
Db 1381 gacccagctgtgtctgtgtgagatctcgagatccgcgaagaaagacccgaagcccgcaa 1440  
QY 1441 tgcgcgcccggtacaaagcccgctgtcgagagggagacacgcgagagctcgtgtatct 1500  
Db 1441 tgcgcgcccggtacaaagcccgctgtcgagagggagacacgcgagagctcgtgtatct 1500  
QY 1501 cgcgcgctctgaaacagcgctgacagcgtgtgttcaaccggaagagatccgcgtgtc 1560  
Db 1501 cgcgcgctctgaaacagcgctgacagcgtgtgttcaaccggaagagatccgcgtgtc 1560  
QY 1561 ccaacacccaccttaagcgaagcgcaagctgtggccctgttcaagtctctgtcgaagaagc 1620  
Db 1561 ccaacacccaccttaagcgaagcgcaagctgtggccctgttcaagtctctgtcgaagaagc 1620  
QY 1621 ggcgcgtgatalccagcctctcaagcgtgacatcgctgcgcggaacccctcccgacacat 1680  
Db 1621 ggcgcgtgatalccagcctctcaagcgtgacatcgctgcgcggaacccctcccgacacat 1680  
QY 1681 ctaccgcgagacgcgcaatctgcgcgcgcgcgaacgctgtctgttggccaatgccaagctca 1740  
Db 1681 ctaccgcgagacgcgcaatctgcgcgcgcgcgaacgctgtctgttggccaatgccaagctca 1740  
QY 1741 gagacacaaagagatcacatcgccagcgaagccgaagccctcgatatacgacagagata 1800  
Db 1741 gagacacaaagagatcacatcgccagcgaagccgaagccctcgatatacgacagagata 1800  
QY 1801 tcaaccttaaccatggaagcttatctcgagaagcgcatcccgccaaagaaacagagc 1860  
Db 1801 tcaaccttaaccatggaagcttatctcgagaagcgcatcccgccaaagaaacagagc 1860  
QY 1861 gttctacacgaatggtgcgtgcgcgaacccgttctgggggaatggagcgtgtatcgagatg 1920  
Db 1861 gttctacacgaatggtgcgtgcgcgaacccgttctgggggaatggagcgtgtatcgagatg 1920



QY 1921 gggccgcatctgacatcgccgcatccggtcgtcatctgtcttgagatgtgacaagaatgc 1980  
|||||  
Db 1921 gggccgcatctgacatcgccgcatccggtcgtcatctgtcttgagatgtgacaagaatgc 1980  
QY 1981 catctgtcgatctgtcgcatctgagaccgccaagcgttcaagcggtgatactgtgctgagcc 2040  
|||||  
Db 1981 catctgtcgatctgtcgcatctgagaccgccaagcgttcaagcggtgatactgtgctgagcc 2040  
QY 2041 catccaatgtgacatgcttcccaaggggacataaagccatacatgtataagatgtgacaagcgga 2100  
|||||  
Db 2041 catccaatgtgacatgcttcccaaggggacataaagccatacatgtataagatgtgacaagcgga 2100  
QY 2101 gcaagcacaagccattccacagttgaggttgagagatgttccggttccacagtgaagcgct 2160  
|||||  
Db 2101 gcaagcacaagccattccacagttgaggttgagagatgttccggttccacagtgaagcgct 2160  
QY 2161 gatgtcttccacgccaagcagctgtacatacgacaagcaagcggtgtctgcgcgacagac 2220  
|||||  
Db 2161 gatgtcttccacgccaagcagctgtacatacgacaagcaagcggtgtctgcgcgacagac 2220  
QY 2221 cgggttcgcccgcaccccgcaagaacgaggtaaatgtgcgatttccgcaaaaaaacggtg 2280  
|||||  
Db 2221 cgggttcgcccgcaccccgcaagaacgaggtaaatgtgcgatttccgcaaaaaaacggtg 2280  
QY 2281 caaatgtgtgcaaaatcacacatccagtttcatctcgtgaaacccgctcgtccaacatgaacgag 2340  
|||||  
Db 2281 caaatgtgtgcaaaatcacacatccagtttcatctcgtgaaacccgctcgtccaacatgaacgag 2340  
QY 2341 caagccatcatccaaagccccaagaaacggtgtgcgagctacagatgagcgaatgttctcgtg 2400  
|||||  
Db 2341 caagccatcatccaaagccccaagaaacggtgtgcgagctacagatgagcgaatgttctcgtg 2400  
QY 2401 ctctctaaagctgtgcaagggccttgcaacagttgatattcaacggtgagatgtgcaggtctcttggc 2460  
|||||  
Db 2401 ctctctaaagctgtgcaagggccttgcaacagttgatattcaacggtgagatgtgcaggtctcttggc 2460  
QY 2461 tctccccaagagagcacctcaggggtgagcgagctagccggtctcaggttcaacagtgtaaatcgt 2520  
|||||  
Db 2461 tctccccaagagagcacctcaggggtgagcgagctagccggtctcaggttcaacagtgtaaatcgt 2520  
QY 2521 ctgagagagcgttgcggggcttatagtgttgcgtgctggtlcaacggttgcgcatcggaatgagcat 2580  
|||||  
Db 2521 ctgagagagcgttgcggggcttatagtgttgcgtgctggtlcaacggttgcgcatcggaatgagcat 2580  
QY 2581 acgattgcttcttaagcagtgcaatccctgagagcttcaacgttggaataaataacgctccaaaaa 2640  
|||||  
Db 2581 acgattgcttcttaagcagtgcaatccctgagagcttcaacgttggaataaataacgctccaaaaa 2640  
QY 2641 agccctgaaccaaatcttgcgaaaaatctgcttgaagaatttgccttcaaaaaacttggaacg 2700  
|||||  
Db 2641 agccctgaaccaaatcttgcgaaaaatctgcttgaagaatttgccttcaaaaaacttggaacg 2700  
QY 2701 agatactgacaagatccctctacgagtgctgtagagtaatgcaatgcaaaacgcat 2760  
|||||  
Db 2701 agatactgacaagatccctctacgagtgctgtagagtaatgcaatgcaaaacgcat 2760  
QY 2761 tttttgcccagtagagagtaatgagatgttattttttggagaaatttgcctcagtaga 2820  
|||||  
Db 2761 tttttgcccagtagagagtaatgagatgttattttttggagaaatttgcctcagtaga 2820  
QY 2821 gtaacgctgtgttaaatcttgcctgagtgaggttcaaatgacacacgagctgcgcttg 2880  
|||||  
Db 2821 gtaacgctgtgttaaatcttgcctgagtgaggttcaaatgacacacgagctgcgcttg 2880  
QY 2881 gtcgtatctgacttgcgcccgcaattggtgacacttgcagggcatcccccctgaaacttctgcgtg 2940  
|||||  
Db 2881 gtcgtatctgacttgcgcccgcaattggtgacacttgcagggcatcccccctgaaacttctgcgtg 2940  
QY 2941 atgacacatttcgaaggaatgtgtcgaaatcatagaaatttgtgtgaggtgcgttagagcg 3000  
|||||  
Db 2941 atgacacatttcgaaggaatgtgtgtcgaaatcatagaaatttgtgtgaggtgcgttagagcg 3000  
QY 3001 ctctgacaggggtgctgcgcgagatctctgtctcaaggtagggcgacaatgagaggtg 3060  
|||||

Db 3001 ctctgacaggggtgctgcgcgagatctctgtctcaaggtagggcgacaatgagaggtg 3060  
|||||  
QY 3061 ttaattgccccctgatatctgcctctgcgtgagcatgtgtgtcacaatccctgcgggaatatg 3120  
|||||  
Db 3061 ttaattgccccctgatatctgcctctgcgtgagcatgtgtgtcacaatccctgcgggaatatg 3120  
QY 3121 atatccgctagagatatacgaatgatttccgctgcgtgcgtgtgtgcggctgtgcgggc 3180  
|||||  
Db 3121 atatccgctagagatatacgaatgatttccgctgcgtgcgtgtgtgtgcggctgtgtgcgggc 3180  
QY 3181 ttgttcgggtctgtcgggcctgtccctctgtcccgccgctgcctcatttccaacataa 3240  
|||||  
Db 3181 ttgttcgggtctgtcgggcctgtccctctgtcccgccgctgcctcatttccaacataa 3240  
QY 3241 aaatgtgcggaagccctctctgttctatagttctataatgttcaatgacaataatcacata 3300  
|||||  
Db 3241 aaatgtgcggaagccctctctgttctatagttctataatgttcaatgacaataatcacata 3300  
QY 3301 atatacaatagcttatctgccttaaaaggagtaatttggccgcaaaaggagtaattgg 3360  
|||||  
Db 3301 atatacaatagcttatctgccttaaaaggagtaatttggccgcaaaaggagtaattgg 3360  
QY 3361 ccgcacaaggagatgaattggtgcgcaaaaggagtaatttggccgcatctcggttcttac 3420  
|||||  
Db 3361 ccgcacaaggagatgaattggtgcgcaaaaggagtaatttggccgcatctcggttcttac 3420  
QY 3421 atgtggaggaatccctcttaacatcttccccaatgggaagaacaaacaagtgtgcgag 3480  
|||||  
Db 3421 atgtggaggaatccctcttaacatcttccccaatgggaagaacaaacaagtgtgcgag 3480  
QY 3481 accggccttcgaaccaacaacaaatgtgtctccctcccgccgaggtgaggaagggtcta 3540  
|||||  
Db 3481 accggccttcgaaccaacaacaaatgtgtctccctcccgccgaggtgaggaagggtcta 3540  
QY 3541 tgcgcaatccgcccgccttgcaaggtgcctcaaggtcattatgataagcaactgtcg 3600  
|||||  
Db 3541 tgcgcaatccgcccgccttgcaaggtgcctcaaggtcattatgataagcaactgtcg 3600  
QY 3601 gtcgcaatccgcccgccttgcaaggtgcctcaaggtcattatgataagcaactgtcg 3660  
|||||  
Db 3601 gtcgcaatccgcccgccttgcaaggtgcctcaaggtcattatgataagcaactgtcg 3660  
QY 3661 acggcatgaaaaaacatacggtgagagcctgcgacccgctgttgaggaagctagccgctg 3720  
|||||  
Db 3661 acggcatgaaaaaacatacggtgagagcctgcgacccgctgttgaggaagctagccgctg 3720  
QY 3721 cgtgttgaaccaatgataccctgcgaagatgatactgtgacagtcggtgcgtgtgcgaatg 3780  
|||||  
Db 3721 cgtgttgaaccaatgataccctgcgaagatgatactgtgacagtcggtgcgtgtgcgaatg 3780  
QY 3781 aggcggaataagatacccgcaagagcgcaagcggaacctccagtgacgttgagacctcc 3840  
|||||  
Db 3781 aggcggaataagatacccgcaagagcgcaagcggaacctccagtgacgttgagacctcc 3840  
QY 3841 ggaatgacatccgtctgatactgcggtgaggtgcgaacacacacggtgcattccgaacgctaca 3900  
|||||  
Db 3841 ggaatgacatccgtctgatactgcggtgaggtgcgaacacacacggtgcattccgaacgctaca 3900  
QY 3901 cgtatctccatctcgttagtaagatattccgtgtcgtgttccaggaagcttccatagctcg 3960  
|||||  
Db 3901 cgtatctccatctcgttagtaagatattccgtgtcgtgttccaggaagcttccatagctcg 3960  
QY 3961 ccaatcttgatccggaatgagcggaacaccttaagtcggtcccgaggtgcgggcctccttg 4020  
|||||  
Db 3961 ccaatcttgatccggaatgagcggaacaccttaagtcggtcccgaggtgcgggcctccttg 4020  
QY 4021 gagtgcggaaggaaagtgtgtgttggaacgaagcttaaacgaatttgcctccaacctg 4080  
|||||  
Db 4021 gagtgcggaaggaaagtgtgtgttggaacgaagcttaaacgaatttgcctccaacctg 4080  
QY 4081 cacttgagagatcaaacattatctgcgtctgacattgacggcaagcggaacgaagtgtg 4140  
|||||



Db	4081	cactgtatagatcaacattatctgcgtctctacattgacgccaagcgccaagattg	4140
QY	4141	gccgtgacgttgacaagtctgactatagctgtggaagtgaagaacacccacgtcgcga	4200
Db	4141	gcccgtacgttgacaagtctgactatagctgtggaagtgaagaacacccacgtcgcga	4200
QY	4201	ggcgagacgtcgaggtttccaaagtcgtcgcagatgctcgtcgcagagggcgacgga	4260
Db	4201	ggcgagacgtcgaggtttccaaagtcgtcgcagatgctcgtcgcagagggcgacgga	4260
QY	4261	cgatagacccctcctccccaagacggcggtgatacctaagctccacgtctgctggagc	4320
Db	4261	cgatagacccctcctccccaagacggcggtgatacctaagctccacgtctgctggagc	4320
QY	4321	tgaacgctctgtctgagacaagaacagacatctgctcgcgtccagactccggcgt	4380
Db	4321	tgaacgctctgtctgagacaagaacagacatctgctcgcgtccagactccggcgt	4380
QY	4381	tcctcgagagagcgctgctctgagacgtgcaaacatcgaaaaactgttttagatt	4440
Db	4381	tcctcgagagagcgctgctctgagacgtgcaaacatcgaaaaactgttttagatt	4440
QY	4441	tcctcgcaaaagttaggaaggtttgagttttgaggtatttcaccgcaatagtgttaag	4500
Db	4441	tcctcgcaaaagttaggaaggtttgagttttgaggtatttcaccgcaatagtgttaag	4500
QY	4501	acttcgtgaaacgatctgacataatagcgttaagactatgaatacacggtcgagacgc	4560
Db	4501	acttcgtgaaacgatctgacataatagcgttaagactatgaatacacggtcgagacgc	4560
QY	4561	tgcataaagcaacggtgtgtgacgacccgcaacacacacacacacacacacacacac	4620
Db	4561	tgcataaagcaacggtgtgtgacgacccgcaacacacacacacacacacacacacac	4620
QY	4621	ttccgttaaaaaagatgatactcggtgacgtggtatagatccctcgagattgcaagat	4680
Db	4621	ttccgttaaaaaagatgatactcggtgacgtggtatagatccctcgagattgcaagat	4680
QY	4681	gtttccctccatttcaaaagaatacacccgaaacacacacacacacacacacacacac	4740
Db	4681	gtttccctccatttcaaaagaatacacccgaaacacacacacacacacacacacacac	4740
QY	4741	tatatgaacacatgaatactcctcagaatacagcgcatlaagcgtgaagtctcgactt	4800
Db	4741	tatatgaacacatgaatactcctcagaatacagcgcatlaagcgtgaagtctcgactt	4800
QY	4801	acgagatgcttatctgatactgacagggagatcgcgcacaaatggtgcgcagacgc	4860
Db	4801	acgagatgcttatctgatactgacagggagatcgcgcacaaatggtgcgcagacgc	4860
QY	4861	tccttcaatttcaacgagatgagaggaagacgccccctcctcaaaaaaagaatgtg	4920
Db	4861	tccttcaatttcaacgagatgagaggaagacgccccctcctcaaaaaaagaatgtg	4920
QY	4921	gaagatactctgatactctgaggtctcagagacgttgccttaaaacctgaatcagcat	4980
Db	4921	gaagatactctgatactctgaggtctcagagacgttgccttaaaacctgaatcagcat	4980
QY	4981	ggcagctgataagaataataatagcccaatagagcgccatttccatccatcaata	5040
Db	4981	ggcagctgataagaataataatagcccaatagagcgccatttccatccatcaata	5040
QY	5041	gctacatctgatacaatacaagtattgatatccatcaatgagagaattacaatgta	5100
Db	5041	gctacatctgatacaatacaagtattgatatccatcaatgagagaattacaatgta	5100
QY	5101	tcacagatcatcagacattgttttctgatttctaagtgctaaactaactcgcgtg	5160
Db	5101	tcacagatcatcagacattgttttctgatttctaagtgctaaactaactcgcgtg	5160
QY	5161	cactaaagaatgttactatctgacatcatcaccctgggtttcagacgcaataatcag	5220
Db	5161	cactaaagaatgttactatctgacatcatcaccctgggtttcagacgcaataatcag	5220
QY	5221	cttgttcgaagccggtttttttccagagagagacacgtttttgattgtgtataaactg	5280
Db	5221	cttgttcgaagccggtttttttccagagagagacacgtttttgattgtgtataaactg	5280
QY	5281	ctagatctccatgattgtatacatcacagcatcccggtgggacacacacacacacacac	5340
Db	5281	ctagatctccatgattgtatacatcacagcatcccggtgggacacacacacacacacac	5340
QY	5341	agatgataatattagcccaagatctgagcggaataacgagattgagccagacggccg	5400
Db	5341	agatgataatattagcccaagatctgagcggaataacgagattgagccagacggccg	5400
QY	5401	cactttaagcgcggaagccataacgagatcccaaaacacatagaagccagattgagaaa	5460
Db	5401	cactttaagcgcggaagccataacgagatcccaaaacacatagaagccagattgagaaa	5460
QY	5461	taggaaacgtgcccctcgagcgctgagctgagctagctagcagatccctttaaataatg	5520
Db	5461	taggaaacgtgcccctcgagcgctgagctgagctagctagcagatccctttaaataatg	5520
QY	5521	tacagaagaagaacatatacagcgccagacaccccgacagcagatccctgttgcattgt	5580
Db	5521	tacagaagaagaacatatacagcgccagacaccccgacagcagatccctgttgcattgt	5580
QY	5581	accgctgcatctccagaaatcctaaacaaagtccgagacccgacacggtcgtctccagc	5640
Db	5581	accgctgcatctccagaaatcctaaacaaagtccgagacccgacacggtcgtctccagc	5640
QY	5641	atcagttctgtttagagagatgtgagacagcgtacagatcttgacacacacacacacac	5700
Db	5641	atcagttctgtttagagagatgtgagacagcgtacagatcttgacacacacacacacac	5700
QY	5701	atcagatcagatcgagac	5760
Db	5701	atcagatcagatcgagac	5760
QY	5761	agaaaaagacggtgtgtcttagaatcgagagctacagctctgacatgtgcataatg	5820
Db	5761	agaaaaagacggtgtgtcttagaatcgagagctacagctctgacatgtgcataatg	5820
QY	5821	aagtggcgcgctgacatcagatctctcctcgaatacctccacacacacacacacacac	5880
Db	5821	aagtggcgcgctgacatcagatctctcctcgaatacctccacacacacacacacacac	5880
QY	5881	gcaatcgcacccctgggttttcccgatattcttcaagatcccaagacgaagctcccat	5940
Db	5881	gcaatcgcacccctgggttttcccgatattcttcaagatcccaagacgaagctcccat	5940
QY	5941	atccacccaagagtgctgaataatggtggtcgcgaacacagataatccgatatctatc	6000
Db	5941	atccacccaagagtgctgaataatggtggtcgcgaacacagataatccgatatctatc	6000
QY	6001	taagtcaattctgacgtcgctctcgaataatggtggtcgcgaacacagataatccg	6060
Db	6001	taagtcaattctgacgtcgctctcgaataatggtggtcgcgaacacagataatccg	6060
QY	6061	tttaagagatttgaactgtgattatcttccctgaaagagacagattactactactacta	6120
Db	6061	tttaagagatttgaactgtgattatcttccctgaaagagacagattactactactacta	6120
QY	6121	ttgttgcgatttcaaacctcaaaacagatggtattgtaatgtaagacaggtgaatcac	6180
Db	6121	ttgttgcgatttcaaacctcaaaacagatggtattgtaatgtaagacaggtgaatcac	6180
QY	6181	ccgctgataatccacaaaggtgagtcctgtagacagacatcccaaggaataacggttta	6240
Db	6181	ccgctgataatccacaaaggtgagtcctgtagacagacatcccaaggaataacggttta	6240
QY	6241	atccatctccatgagatcaacaatagatggtgttcaggtcccgacatctggtgtgaggt	6300
Db	6241	atccatctccatgagatcaacaatagatggtgttcaggtcccgacatctggtgtgaggt	6300







Db 8461 cggagacacgaagctgcgcgcgcgtatgtgtgacacgcgcga 8509

RESULT 2

ID AAS18307 standard; DNA; 8509 BP.

XX AAS18307;

XX 12-MAR-2002 (first entry)

DE DNA sequence of Ketogulonigenium endogenous plasmid pADM291.

XX Cloning vector: Ketogulonigenium replicon; endogenous plasmid;

KW transformed host cell; Escherichia coli; pADM291; circular; cyclic; ds.

XX Ketogulonigenium sp. strain ADM291-19.

OS WO200177347-A2.

PN 18-OCT-2001.

PD 05-APR-2001; 2001WO-US11059.

XX 05-APR-2001; 2001WO-US11059.

XX 05-APR-2001; 2000US-194625P.

PR (ARCH ) ARCHER-DANIELS MIDLAND CO.

XX (DELI/) D'ELIA J.

PA D'Elia J;

PI WPI: 2002-049150/06.

DR Novel nucleic acid vector comprising Ketogulonigenium replicon found on

XX a specific deposited endogenous plasmid, useful for producing

PT polypeptides and/or transcripts by culturing host cells transformed

PT with vector -

XX Example 7: Fig 2: 66bp; English.

XX The present invention relates to the isolation of vectors comprising

CC a Ketogulonigenium replicon found on the endogenous plasmid, pADM291.

CC The invention also describes methods of transforming host cells with

CC the vectors and producing polypeptides and/or antisense transcripts by

CC culturing the transformed host cells. The vectors are useful for

CC transforming a host cell by conjugation or electroporation.

CC The vectors which have a replicon functional in both Ketogulonigenium

CC and Escherichia coli, enable the cloning of certain genes of

CC Ketogulonigenium in E.coli as the latter is an efficient host for

CC amplification of vector DNA. The present DNA sequence represents the

CC Ketogulonigenium endogenous plasmid pADM291.

XX Sequence 8509 BP; 2065 A; 2239 C; 2345 G; 1860 T; 0 other;

SO

Query Match 100.0%; Score 8509; DB 24; Length 8509;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 8509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggtagacgactgtgctgtgttgaagaaacgaaagcccgccacagttgtggt 60

Db 1 tggtagacgactgtgctgtgttgaagaaacgaaagcccgccacagttgtggt 60

QY 61 agagcgtcgatattgcaagacgctgctgtgacatcgaaactcgtgtctaactcgc 120

Db 61 agagcgtcgatattgcaagacgctgctgtgacatcgaaactcgtgtctaactcgc 120

QY 121 atcgtcgtatcatagcctgctgcgattgcaagccgacctatgtgtcgaagc 180

Db 121 atcgtcgtatcatagcctgctgcgattgcaagccgacctatgtgtcgaagc 180

QY 161 ccttaagcgatgcgcgacttgcccatgacgactcgtgtgttcaacatcgccgat 240

Db 161 ccttaagcgatgcgcgacttgcccatgacgactcgtgtgttcaacatcgccgat 240

Db 181 ccttaagcgatgcgcgacttgcccatgacgactcgtgtgttcaacatcgccgat 240

QY 241 taacgaactggaatgatactctagaagcggttaaccgagacatctgttaacgcataaag 300

Db 241 taacgaactggaatgatactctagaagcggttaaccgagacatctgttaacgcataaag 300

QY 301 ctgagcgatgaataaataagcgacacatctgccttgcgcgcgccttctcaagcgtgct 360

Db 301 ctgagcgatgaataaataagcgacacatctgccttgcgcgcgccttctcaagcgtgct 360

QY 361 gctctcgtcgtcaatggaacgacgacagatcgcgttccatatacaggtctgtatatacc 420

Db 361 gctctcgtcgtcaatggaacgacgacagatcgcgttccatatacaggtctgtatatacc 420

QY 421 gccgccttgagccatgcacgacacatcgcatcaagccgcgcgtgtatcgtctgtat 480

Db 421 gccgccttgagccatgcacgacacatcgcatcaagccgcgcgtgtatcgtctgtat 480

QY 481 cggcgatagaagcgccacgctcgcgtcgtccatgtggtcttgatagctcgggaac 540

Db 481 cggcgatagaagcgccacgctcgcgtcgtccatgtggtcttgatagctcgggaac 540

QY 541 atcgttgcgtcgtatcgtcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600

Db 541 atcgttgcgtcgtatcgtcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600

QY 601 ctcctcgtgaagcgagcgagcagatcgaaggtgatcaacatcgtcatatctcttc 660

Db 601 ctcctcgtgaagcgagcgagcagatcgaaggtgatcaacatcgtcatatctcttc 660

QY 661 tttcctcgtggtatgacgc 720

Db 661 tttcctcgtggtatgacgc 720

QY 721 cagcctcgtatcaggtggtggaacacgaagcgagcgacatagatctgcaccaag 780

Db 721 cagcctcgtatcaggtggtggaacacgaagcgagcgacatagatctgcaccaag 780

QY 781 ttgccaagcgatttccgc 840

Db 781 ttgccaagcgatttccgc 840

QY 841 cgaacacatgcacagacatctcgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 900

Db 841 cgaacacatgcacagacatctcgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 900

QY 901 tggcggaatccgaataaaccgcatcttcatgacgaatataactacacatcttgat 960

Db 901 tggcggaatccgaataaaccgcatcttcatgacgaatataactacacatcttgat 960

QY 961 attcaatggaagcgaggttcaaatcaacgcccccaacgcgaactgtatctctac 1020

Db 961 attcaatggaagcgaggttcaaatcaacgcccccaacgcgaactgtatctctac 1020

QY 1021 tcaacgcgccttctgaatagaagcttgcacatgataacccgcgcgcgcgcgcgcgcgcgcgc 1080

Db 1021 tcaacgcgccttctgaatagaagcttgcacatgataacccgcgcgcgcgcgcgcgcgcgcgc 1080

QY 1081 aggcgaatccgc 1140

Db 1081 aggcgaatccgc 1140

QY 1141 tactgaagaacatgcaaggaattgcgccggaatgagcaacgaacgaaccccaagcc 1200

Db 1141 tactgaagaacatgcaaggaattgcgccggaatgagcaacgaacgaaccccaagcc 1200

QY 1201 ggcctggaacaaagcgcgctgttgcggaagaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1260

Db 1201 ggcctggaacaaagcgcgctgttgcggaagaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1260

QY 1261 cctgacgcgttcacacgc 1320

Db 1261 cctgacgcgttcacacgc 1320



[illegible]

QY	2401	ctcAtaggctgcgaagggccctgcgaagaatgatttcaacggtgagatgtgcaaggtcttttggc	2460
Db	2401	ctcAtaaggctgcgaagggccctgcgaagaatgatttcaacggtgagatgtgcaaggtcttttggc	2460
QY	2461	tctcccgcaagaagccaacttcagaggtgagacgagctctagctttaggtttcaacgtgaaatcgc	2520
Db	2461	tctcccgcaagaagccaacttcagaggtgagacgagctctagctttaggtttcaacgtgaaatcgc	2520
QY	2521	ctgaggaagcgtctgcggggtctcAtgttttggctctgtgtcAatcgtctgcacatcagatgtgagcat	2580
Db	2521	ctgaggaagcgtctgcggggtctcAtgttttggctctgtgtcAatcgtctgcacatcagatgtgagcat	2580
QY	2581	acgaAtgctcttaagcgagtcgaaatccctgtagagcttcaacgttgggaaaaatacgtctccaaaa	2640
Db	2581	acgaAtgctcttaagcgagtcgaaatccctgtagagcttcaacgttgggaaaaatacgtctccaaaa	2640
QY	2641	agccctgcgcaaatcttggaaaaatttgtttttaaagtttgtttttaaanaacttgggaacg	2700
Db	2641	agccctgcgcaaatcttggaaaaatttgtttttaaagtttgtttttaaanaacttgggaacg	2700
QY	2701	agataAtgacagagatccctcttaacgAtgtctgtcAagagtaatcgaAtgagcaaaaacgcat	2760
Db	2701	agataAtgacagagatccctcttaacgAtgtctgtcAagagtaatcgaAtgagcaaaaacgcat	2760
QY	2761	tttttgcgcccaatgagagtaatgtgataatttttgggagatlttgtcttcagtagga	2820
Db	2761	tttttgcgcccaatgagagtaatgtgataatttttgggagatlttgtcttcagtagga	2820
QY	2821	gtAaagcggttggttAaatttgtctgtgatttggctgcgttccaaatcgacaacgagctgcgtltg	2880
Db	2821	gtAaagcggttggttAaatttgtctgtgatttggctgcgttccaaatcgacaacgagctgcgtltg	2880
QY	2881	gtcgcAttcgcgaAtctcccccgaatattgggcactgtgagacgaatcccccgtAaacttctggcg	2940
Db	2881	gtcgcAttcgcgaAtctcccccgaatattgggcactgtgagacgaatcccccgtAaacttctggcg	2940
QY	2941	atgaccattctcgaaagcgaatgggtctgaaatctcaatagaatlttgtgtgaggtgtgcgtlaacg	3000
Db	2941	atgaccattctcgaaagcgaatgggtctgaaatctcaatagaatlttgtgtgaggtgtgcgtlaacg	3000
QY	3001	ctctgaacggggtgtgtctgcgcggagatctctctgtctcAgtatgggcgcgaacatvggagaggtg	3060
Db	3001	ctctgaacggggtgtgtctgcgcggagatctctctgtctcAgtatgggcgcgaacatvggagaggtg	3060
QY	3061	ttagttgcgccctgtcatcgcctctctgcgttggcgcaatgggtgcAatccctgcgccgagacatgt	3120
Db	3061	ttagttgcgccctgtcatcgcctctctgcgttggcgcaatgggtgcAatccctgcgccgagacatgt	3120
QY	3121	atattccgctagaggaattactgtatagtttctgtcctgtcgcgttgcgggctgtgcgggtctgcgggc	3180
Db	3121	atattccgctagaggaattactgtatagtttctgtcctgtcgcgttgcgggctgtgcgggtctgcgggc	3180
QY	3181	tgtctgcgggtctgtgcgggctgtctccctctgtcccgccgtccctcaattttcacaatcaa	3240
Db	3181	tgtctgcgggtctgtgcgggctgtctccctctgtcccgccgtccctcaattttcacaatcaa	3240
QY	3241	aaaatvgggcgaaagccctctctgttctatagtttctatagtttcataagttcacaagaaattacacata	3300
Db	3241	aaaatvgggcgaaagccctctctgttctatagtttctatagtttcataagttcacaagaaattacacata	3300
QY	3301	atctatcaatagcttctatcgtctttaaaggagtaatttggccgcgcaaaaaggaggtatctgg	3360
Db	3301	atctatcaatagcttctatcgtctttaaaggaggtatcttggccgcgcaaaaaggaggtatctgg	3360
QY	3361	ccgcaaaaaggagtaatttgggcgcgcaaaaaggagtaatttgggcgcgatactgcgttgtttac	3420
Db	3361	ccgcaaaaaggagtaatttgggcgcgcaaaaaggagtaatttgggcgcgatactgcgttgtttac	3420
QY	3421	atcggggaaggaatccctcttaatcatcttcccccatactgggagaaagacacaaagttggccgag	3480
Db	3421	atcggggaaggaatccctcttaatcatcttcccccatactgggagaaagacacaaagttggccgag	3480
QY	3481	accgggctctgcacagcaaaaactgtgtctccctgcgcgaaggtgtgcgaggggtctata	3540



[illegible]

Db	4561	tgcaaaagcaacgggtgtgtgycgacgcgaacccatcactcgggcgtctaaagagcgtt	4620
QY	4621	ttccggtaaaaaagatcaatccttggggcatgtgtatagatctctgcagaaatgtcacagat	4680
Db	4621	ttccggtaaaaaagatcaatccttggggcatgtgtatagatctctgcagaaatgtcacagat	4680
QY	4681	gtttccctcccatltaaaagaatacacccgaacaactaaacgcgaagtatagtgaagc	4740
Db	4681	gtttccctcccatltaaaagaatacacccgaacaactaaacgcgaagtatagtgaagc	4740
QY	4741	tgatgaacacatgaaatgacctgaagaatcaagcgaattgaagcgtgaagtttcgacatt	4800
Db	4741	tgatgaacacatgaaatgacctgaagaatcaagcgaattgaagcgtgaagtttcgacatt	4800
QY	4801	acgcgaatgcttatcatgtgtccaggaggatcgcgcgaacaatgtgcgcgacatgtccagacg	4860
Db	4801	acgcgaatgcttatcatgtgtccaggaggatcgcgcgaacaatgtgcgcgacatgtccagacg	4860
QY	4861	tcttcaatttcaatccgatatgagagaggaagacgcccctcaaaaacaaagatgtgtg	4920
Db	4861	tcttcaatttcaatccgatatgagagaggaagacgcccctcaaaaacaaagatgtgtg	4920
QY	4921	gaagatattctgacctcgtggtcttcaggagccttgcctttaaacctgaatcagatctta	4980
Db	4921	gaagatattctgacctcgtggtcttcaggagccttgcctttaaacctgaatcagatctta	4980
QY	4981	gcgatgctgtataagaagttaataatagcccacaataagagcgccatttccatctcacata	5040
Db	4981	gcgatgctgtataagaagttaataatagcccacaataagagcgccatttccatctcacata	5040
QY	5041	gctcatcatgtgatacatalcaagatgtgatatcacaatgtgagagaatattcaagtta	5100
Db	5041	gctcatcatgtgatacatalcaagatgtgatatcacaatgtgagagaatattcaagtta	5100
QY	5101	tcaacagatcatcaacgcatgtttgtttgatatcttaagtgtgtaacaaactatcgtctg	5160
Db	5101	tcaacagatcatcaacgcatgtttgtttgatatcttaagtgtgtaacaaactatcgtctg	5160
QY	5161	cccctaaagaagatgtgactaattgcagatctcaaccttgggttcagacgcgataattaag	5220
Db	5161	cccctaaagaagatgtgactaattgcagatctcaaccttgggttcagacgcgataattaag	5220
QY	5221	cttgtgcgaagcgggttttttgcagagagagagacggttttgatggtgtataaactg	5280
Db	5221	cttgtgcgaagcgggttttttgcagagagagagacggttttgatggtgtataaactg	5280
QY	5281	ctaacgtactcatgtatgtgatacacacacgacatcatccggggcaacaacacctctatgtga	5340
Db	5281	ctaacgtactcatgtatgtgatacacacacgacatcatccggggcaacaacacctctatgtga	5340
QY	5341	agatgatatattatgcccagagaatgtaactgctggcgggggaaaaaacgcgatgtgaagccgcg	5400
Db	5341	agatgatatattatgcccagagaatgtaactgctggcgggggaaaaaacgcgatgtgaagccgcg	5400
QY	5401	gaccttaagggcggaagcctataacgcgttaccacgaatacacaactagaaagccagatgtgaagaaa	5460
Db	5401	gaccttaagggcggaagcctataacgcgttaccacgaatacacaactagaaagccagatgtgaagaaa	5460
QY	5461	taagggaacgtgtccctcctgagcgctgtgcgctgctagagatcccttttaataatgtgtag	5520
Db	5461	taagggaacgtgtccctcctgagcgctgtgcgctgctagagatcccttttaataatgtgtag	5520
QY	5521	tataagaagaacagacaattatatacagcgccgacatccgacgcgcagatctctgttatctgt	5580
Db	5521	tataagaagaacagacaattatatacagcgccgacatccgacgcgcagatctctgttatctgt	5580
QY	5581	accgcgtgcacatccagaatctcaaaccaagtcgagagccggaacgcgctcgtcttcacgc	5640
Db	5581	accgcgtgcacatccagaatctcaaaccaagtcgagagccggaacgcgctcgtcttcacgc	5640
QY	5641	atcaagtttctgttagagagatgtgagacgcgtctacagatgtgacacccgtacatgccc	5700
Db	5641	atcaagtttctgttagagagatgtgagacgcgtctacagatgtgacacccgtacatgccc	5700



QY	5701	atcgactcogatcgagcaactactagctcgggtaacagatvgtaactaaagttaacagtttatbga	5760
Db	5701	atcgactcogatcgagcaactactagctcgggtaacagatvgtaactaaagttaacagtttatbga	5760
QY	5761	agaaaaagacggtttggtctctgaatctcgagagcgtacagctctgtaatgctcatnaatltgaaat	5820
Db	5761	agaaaaagacggtttggtctctgaatctcgagagcgtacagctctgtaatgctcatnaatltgaaat	5820
QY	5821	aagtgcgcgctcgatctcagltcagltacatcttcctgcgtacatcccaacaatbgtatcgtttaatg	5880
Db	5821	aagtgcgcgctcgatctcagltcagltacatcttcctgcgtacatcccaacaatbgtatcgtttaatg	5880
QY	5881	gcaattcgcaacctbgtggtttctctcgatattactttagatccaagaagccagaagctcccat	5940
Db	5881	gcaattcgcaacctbgtggtttctctcgatattactttagatccaagaagccagaagctcccat	5940
QY	5941	atccaccaagaagtgctgaataatactgggtctcgagaactcagaataatctacglatctatctc	6000
Db	5941	atccaccaagaagtgctgaataatactgggtctcgagaactcagaataatctacglatctatctc	6000
QY	6001	tatcgactctcgcactbgcgctctccgtaaaatgataccaagataagatgatacattacatctcga	6060
Db	6001	tatcgactctcgcactbgcgctctccgtaaaatgataccaagataagatgatacattacatctcga	6060
QY	6061	tttagagatlttagctgtgtatctctctccctgaagaagagacagatattactctcaacta	6120
Db	6061	tttagagatlttagctgtgtatctctctccctgaagaagagacagatattactctcaacta	6120
QY	6121	ttgtgcgcggtattccaacctcaaaaacagatgataattggaatgtgcagagaaagtggaatcac	6180
Db	6121	ttgtgcgcggtattccaacctcaaaaacagatgataattggaatgtgcagagaaagtggaatcac	6180
QY	6181	ccgcgtgatattccaaccaaggttgagatccctgtagatcagactctcaagaggtaaacgtttta	6240
Db	6181	ccgcgtgatattccaaccaaggttgagatccctgtagatcagactctcaagaggtaaacgtttta	6240
QY	6241	atccatctccatgatatccaacatagatagaagtgttcaagctcccgatctcgtgcgatccggt	6300
Db	6241	atccatctccatgatatccaacatagatagaagtgttcaagctcccgatctcgtgcgatccggt	6300
QY	6301	cttagagatgaatctgtccggtctcttgacataacccccgcgtgcgaaccccttcttcaaaag	6360
Db	6301	cttagagatgaatctgtccggtctcttgacataacccccgcgtgcgaaccccttcttcaaaag	6360
QY	6361	aaagtccagcgcgctctcgaaagccgctctttagccggaatgccccggccaaacggaatgaaagc	6420
Db	6361	aaagtccagcgcgctctcgaaagccgctctttagccggaatgccccggccaaacggaatgaaagc	6420
QY	6421	gaacaagacgaagcgcctactgatactgtctctgcgcctccggaaagctgcagccgacaatcc	6480
Db	6421	gaacaagacgaagcgcctactgatactgtctctgcgcctccggaaagctgcagccgacaatcc	6480
QY	6481	ggaaaaaagaaaaatacatcagtcagatattttgatatagagacaaatccctttcttata	6540
Db	6481	ggaaaaaagaaaaatacatcagtcagatattttgatatagagacaaatccctttcttata	6540
QY	6541	atatatacaacaagatatactgagcatgcgcgcgcgtgcgtatccctcatcttgataccaatccaaatcc	6600
Db	6541	atatatacaacaagatatactgagcatgcgcgcgcgtgcgtatccctcatcttgataccaatccaaatcc	6600
QY	6601	tgaagaagctgactatagccctactagcatatcgcgacacatctggtgcgcgtataatgataagatgaac	6660
Db	6601	tgaagaagctgactatagccctactagcatatcgcgacacatctggtgcgcgtataatgataagatgaac	6660
QY	6661	cgcacactacacgtgcgcgcgcgaatccaacgcgcctactgctctgtctgcgcgcgaagaagctgt	6720
Db	6661	cgcacactacacgtgcgcgcgcgaatccaacgcgcctactgctctgtctgcgcgcgaagaagctgt	6720
QY	6721	gcgcgtgttcgaactctctgtttgaaaggtatccgatacgaacagcttcaatgcttctbgtgcggtg	6780
Db	6721	gcgcgtgttcgaactctctgtttgaaaggtatccgatacgaacagcttcaatgcttctbgtgcggtg	6780

[illegible]



```

|||||
Db 7861 atccgcatcaccctgctgtagcgccgcatctcccaagtaacgaagctgtcagtcacatg 7920
Oy 7921 gccgaagggggagcgctatctcaagtgtgagtgagtgacatcgaaagaatgtctggaatg 7980
Db 7921 gccgaagggggagcgctatctcaagtgtgagtgagtgacatcgaaagaatgtctggaatg 7980
Oy 7981 caataatgcaagcccggaacctctgtgcgtctgctctatcatcaggcatcagctg 8040
Db 7981 caataatgcaagcccggaacctctgtgcgtctgctctatcatcaggcatcagctg 8040
Oy 8041 aacgcatcgacgtcgatctgcatatcatatcccaagcgagcgctagaanaatgtg 8100
Db 8041 aacgcatcgacgtcgatctgcatatcatatcccaagcgagcgctagaanaatgtg 8100
Oy 8101 ctgcgcgggcgcaactctgcccagcatcgatccatcagaanaatcttggcgcggttaacagt 8160
Db 8101 ctgcgcgggcgcaactctgcccagcatcgatccatcagaanaatcttggcgcggttaacagt 8160
Oy 8161 ttggcgcaactcgcaacttttcgcgagagatgagaataatcgagaaggtcgctgcacatc 8220
Db 8161 ttggcgcaactcgcaacttttcgcgagagatgagaataatcgagaaggtcgctgcacatc 8220
Oy 8221 agtcatgattacccgagcgcggttgattttctaagagagatataatcccgcgccac 8280
Db 8221 agtcatgattacccgagcgcggttgattttctaagagagatataatcccgcgccac 8280
Oy 8281 ctgaactgtgcagatgctacatcctgcgcgagttcgtgcgcgaaacgggaaggtgac 8340
Db 8281 ctgaactgtgcagatgctacatcctgcgcgagttcgtgcgcgaaacgggaaggtgac 8340
Oy 8341 cggcgcggttgatgaggtgtgtcgagcgacgagcgatgaaatgaggtgtgtgtgtcac 8400
Db 8341 cggcgcggttgatgaggtgtgtcgagcgacgagcgatgaaatgaggtgtgtgtgtcac 8400
Oy 8401 gcggaagagctgcgtcaaatctgtctcgacatcgatcttgaaaaaagcgcgctcgaaacccgtg 8460
Db 8401 gcggaagagctgcgtcaaatctgtctcgacatcgatcttgaaaaaagcgcgctcgaaacccgtg 8460
Oy 8461 cgggacaccaggaagcagatcgccgcgcgctatgtgtgtgcacccgagaga 8509
Db 8461 cgggacaccaggaagcagatcgccgcgcgctatgtgtgtgcacccgagaga 8509

RESULT 3
AAS18309
ID AAS18309 standard; DNA: 2517 BP.
XX
AC AAS18309;
XX
DT 12-MAR-2002 (first entry)
XX
DE DNA region of Ketogulonigenium plasmid pADM291 supporting replication.
XX
KW Cloning vector; Ketogulonigenium replicon; endogenous plasmid; pADM291;
XX transformed host cell; Escherichia coli; plasmid vector replication; ds.
XX
OS Ketogulonigenium sp. strain ADM291-19.
XX
PN WO200177347-A2.
XX
PD 18-OCT-2001.
XX
PF 05-APR-2001; 2001WO-US11059.
XX
PR 05-APR-2000; 2000US-194625P.
XX
PA (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX (DELI/) D'ELIA J.
XX
PI D'Elia J;
XX
DR WPI; 2002-049150/06.

```

```

XX
PT Novel nucleic acid vector comprising Ketogulonigenium replicon found on
PT a specific deposited endogenous plasmid, useful for producing
PT polypeptides and/or transcripts by culturing host cells transformed
PT with vector -
XX
PS Claim 4; Fig 4; 66pp; English.
XX
CC The present invention relates to the isolation of vectors comprising
CC Ketogulonigenium replicon found on the endogenous plasmid, pADM291.
CC The invention also describes methods of transforming host cells with
CC the vectors and producing polypeptides and/or antisense transcripts by
CC culturing the transformed host cells. The vectors are useful for
CC transforming a host cell by conjugation or electroporation.
CC The vectors which have a replicon functional in both Ketogulonigenium
CC and Escherichia coli, enable the cloning of certain genes of
CC Ketogulonigenium in E.coli as the latter is an efficient host for
CC amplification of vector DNA. The present DNA sequence represents the
CC region of Ketogulonigenium endogenous plasmid pADM291 that supports
CC plasmid vector replication.
XX
SQ Sequence 2517 BP; 657 A; 589 C; 683 G; 588 T; 0 other:

```

```

Query Match 29.6%; Score 2517; DB 24; Length 2517;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2955 gccgaaggtgcgaattcctagaaatttgatgaggtgcgtacgaggtg 3014
Db 1 gccgaaggtgcgaattcctagaaatttgatgaggtgcgtacgaggtg 60

Oy 3015 ctgcgcgagatctctgtctcagtgagtgagcgacaatgagaggtgtgagttgccccctg 3074
Db 61 ctgcgcgagatctctgtctcagtgagtgagcgacaatgagaggtgtgagttgccccctg 120

Oy 3075 taccgtctctgtgtgcgcatgggttcacacctgcgcgacacatagatctcgtacag 3134
Db 121 taccgtctctgtgtgcgcatgggttcacacctgcgcgacacatagatctcgtacag 180

Oy 3135 gattactgatagtttctgcctgcgttcgagctgtgcggctgtcgggctgtcggctgtc 3194
Db 181 gattactgatagtttctgcctgcgttcgagctgtgcggctgtcgggctgtcggctgtc 240

Oy 3195 gggcgtgccctctgtccgcgctgtccctcacttttcaacaataaagtggcggaac 3254
Db 241 gggcgtgccctctgtccgcgctgtccctcacttttcaacaataaagtggcggaac 300

Oy 3255 cctctgtctatagttcttatagttcatagttcatacgaaataacataatatacagtt 3314
Db 301 cctctgtctatagttcttatagttcatagttcatacgaaataacataatatacagtt 360

Oy 3315 atcgcttaaaaggagtaattggcgcaaaaggagatattggcgccgaaggaggt 3374
Db 361 atcgcttaaaaggagtaattggcgcaaaaggagatattggcgccgaaggaggt 420

Oy 3375 aatggcgccgaaggagtaattggcgcatatcggtgttttaacatggggagatcc 3434
Db 421 aatggcgccgaaggagtaattggcgcatatcggtgttttaacatggggagatcc 480

Oy 3435 cctaatcatctcccatatggaagaacaacaagtggcgagacggcgccctcgac 3494
Db 481 cctaatcatctcccatatggaagaacaacaagtggcgagacggcgccctcgac 540

Oy 3495 cagacaaaactgtgtccctcccgaggtgaggaaggtctatatgcatcccgcc 3554
Db 541 cagacaaaactgtgtccctcccgaggtgaggaaggtctatatgcatcccgcc 600

Oy 3555 cgccgtcagcgctcaagctcatgcatatgataagcacttgcggcgccgcatggt 3614
Db 601 cgccgtcagcgctcaagctcatgcatatgataagcacttgcggcgccgcatggt 660

Oy 3615 gatgatgtgcgcatagaattgcgctggtcggaattcgcgaatcgacgagcatgaanaac 3674

```



[illegible]

Db	1711	caagaataatcacccgaacaacctaaccgcaatlatatgttgtaagcgtgtatgaacaacatg	1800
Qy	4755	aatgaacctcagaataatcagcgcatltagagcglttaagtcttcggaccttaccgatatcttat	4814
Db	1801	aaatgaacctcagaataatcagcgcatltagagcglttaagtcttcggaccttaccgatatcttat	1860
Qy	4815	ctgabyccagggaggaatcgcgcgacaanaatgtgcgcgacatgtgcgagcgctttcaattcat	4874
Db	1861	ctgabyccagggaggaatcgcgcgacaanaatgtgcgcgacatgtgcgagcgctttcaattcat	1920
Qy	4875	caccgatgtagagaggaagaccgccccccctcaaaaagaatgtgtgtgagaatatcttctgat	4934
Db	1921	caccgatgtagagaggaagaccgccccccctcaaaaagaatgtgtgtgagaatatcttctgat	1980
Qy	4935	ccttgagcttcaagagcgcttgcctcttcaaaaaccttaacacgacatcttcagagatgtcgataag	4994
Db	1981	ccttgagcttcaagagcgcttgcctcttcaaaaaccttaacacgacatcttcagagatgtcgataag	2040
Qy	4995	aagtaaatatagccacaatatagagcggcgccattctccatatcaatacagctcatcatgtgat	5054
Db	2041	aagtaaatatagccacaatatagagcggcgccattctccatatcaatacagctcatcatgtgat	2100
Qy	5055	caatatcaagatgtgatattcatcacaatgtgagaagaatttaactgttatacaacgagatcatca	5114
Db	2101	caatatcaagatgtgatattcatcacaatgtgagaagaatttaactgttatacaacgagatcatca	2160
Qy	5115	cagcatgtgttttggatcttcttaagtggtcacaataactctgcctgcccctaaagaagatt	5174
Db	2161	cagcatgtgttttggatcttcttaagtggtcacaataactctgcctgcccctaaagaagatt	2220
Qy	5175	gtactatgtcagtaatctcactcttgggtlcttaacgcgatatcatcagcttltlctgaagcgg	5234
Db	2221	gtactatgtcagtaatctcactcttgggtlcttaacgcgatatcatcagcttltlctgaagcgg	2280
Qy	5235	gtttttttccagagagagacaagctttttgtagtgttaataactcgtacgtatctcatgt	5294
Db	2281	gtttttttccagagagagacaagctttttgtagtgttaataactcgtacgtatctcatgt	2340
Qy	5295	atgtgaataatacacagatatatcccggtgcgaacacaccttctatgtgaagatgtgatatgtg	5354
Db	2341	atgtgaataatacacagatatatcccggtgcgaacacaccttctatgtgaagatgtgatatgtg	2400
Qy	5355	gcccagaagtactcgtgcgaaaaacgcgatatlttaggcacagggcccgcactttaagggcgg	5414
Db	2401	gcccagaagtactcgtgcgaaaaacgcgatatlttaggcacagggcccgcactttaagggcgg	2460
Qy	5415	aagcctataacgagtaccaaaaacactagaagccagatgtgaggaaaaataggagaacgtg	5471
Db	2461	aagcctataacgagtaccaaaaacactagaagccagatgtgaggaaaaataggagaacgtg	2517

RESULT	4
AA518306	
ID	AA518306 standard; DNA; 2112 BP.
XX	
XX	
AC	AA518306;
XX	
DT	12-MAR-2002 (first entry)
XX	
DE	DNA sequence of Ketogulonigenium replicon from plasmid pADM291
XX	
XX	
KM	Cloning vector: Ketogulonigenium replicon; endogenous plasmid;
KM	transformed host cell; Escherichia coli; pADM291; ds.
OS	Ketogulonigenium sp. strain ADM291-19.
XX	
PN	WO200177347-A2.
XX	
PD	18-OCT-2001.
XX	
PF	05-APR-2001; 2001WO-US11059.
XX	
PR	05-APR-2000; 2000US-194625P.
XX	



XX (ARCH ) ARCHER-DANIELS-MIDLAND CO.  
PA (DELI/) D'ELIA J.  
PI D'Elia J;  
XX WPI; 2002-04/9150/06.  
XX Novel nucleic acid vector comprising Ketogulonigenium replicon found on  
PT a specific deposited endogenous plasmid, useful for producing  
PT polypeptides and/or transcripts by culturing host cells transformed  
PT with vector -  
XX Claim 2; Fig 1; 66pp; English.  
XX The present invention relates to the isolation of vectors comprising  
CC a Ketogulonigenium replicon found on the endogenous plasmid, PADM291.  
CC The invention also describes methods of transforming host cells with  
CC the vectors and producing polypeptides and/or antisense transcripts by  
CC culturing the transformed host cells. The vectors are useful for  
CC transforming a host cell by conjugation or electroporation.  
CC The vectors which have a replicon functional in both Ketogulonigenium  
CC and Escherichia coli, enable the cloning of certain genes of  
CC Ketogulonigenium in E.coli as the latter is an efficient host for  
CC amplification of vector DNA. The present DNA sequence represents the  
CC replicon of Ketogulonigenium endogenous plasmid PADM291.  
XX  
XX Sequence 2112 BP; 530 A; 509 C; 605 G; 468 T; 0 other;  
SQ

Query Match 23.6%; Score 2007; DB 24; Length 2112;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2955 ggcacatggttcgaatactatagaattgtgtgaggtcgtagcgctctgacaggagtg 3014  
DB 1 ggcacatggttcgaatactatagaattgtgtgaggtcgtagcgctctgacaggagtg 60  
OY 3015 ctgcgagagatctcgtctcagtgtagggcgacaatgagagtggttaagttgcccctg 3074  
DB 61 ctgcgagagatctcgtctcagtgtagggcgacaatgagagtggttaagttgcccctg 120  
OY 3075 tctcgtctctggtgaggttggttcctcgtccgacatatatgatacttcgcttagag 3134  
DB 121 tctcgtctctggtgaggttggttcctcgtccgacatatatgatacttcgcttagag 180  
OY 3135 gattactgatagttctcgtctcgtcggttgcgtggttcgtcggttcgtcggttcgtc 3194  
DB 181 gattactgatagttctcgtctcgtcggttgcgtggttcgtcggttcgtcggttcgtc 240  
OY 3195 gggccttcctctctcgtccgctcgtcctcacttttcacaatacaaaaatggcgagcg 3254  
DB 241 gggccttcctctctcgtccgctcgtcctcacttttcacaatacaaaaatggcgagcg 300  
OY 3255 cctctgtctatagttctatagttcacaatacaaaaatggcgagcggttgcgttcgtc 3314  
DB 301 cctctgtctatagttctatagttcacaatacaaaaatggcgagcggttgcgttcgtc 360  
OY 3315 atcgcttaaaaggagtaattggcgcaaaaggagtaattggcgcaaaaggagtaatt 3374  
DB 361 atcgcttaaaaggagtaattggcgcaaaaggagtaattggcgcaaaaggagtaatt 420  
OY 3375 aattggcgcaaaaggagtaattggcgcaaaaggagtaattggcgcaaaaggagtaatt 3434  
DB 421 aattggcgcaaaaggagtaattggcgcaaaaggagtaattggcgcaaaaggagtaatt 480  
OY 3435 ccttaatacttcctcccaatgggaaagacacaacaagtggcgcaaaaggagtaattgg 3494  
DB 481 ccttaatacttcctcccaatgggaaagacacaacaagtggcgcaaaaggagtaattgg 540  
OY 3495 cagacaaaactgtgtcctcgtcgaggtggtcgagaggttataatgcgaatccgccc 3554  
DB 541 cagacaaaactgtgtcctcgtcgaggtggtcgagaggttataatgcgaatccgccc 600

OY 3555 ccgctcagcgctcgaagctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 3614  
DB 601 ccgctcagcgctcgaagctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatc 660  
OY 3615 gatgatgtgcgcctgaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3674  
DB 661 gatgatgtgcgcctgaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 720  
OY 3675 catgacgttgagagctgaacccgtcttcgaagagctgagcgtcgtcgtcgtcgtcgtc 3734  
DB 721 catgacgttgagagctgaacccgtcttcgaagagctgagcgtcgtcgtcgtcgtcgtc 780  
OY 3735 gatgacgttcgaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3794  
DB 781 gatgacgttcgaagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 840  
OY 3795 tacgcgcagagcgcaagcgcgcaactcctgtagtcgtgagcgttcgaacgttaccatc 3854  
DB 841 tacgcgcagagcgcaagcgcgcaactcctgtagtcgtgagcgttcgaacgttaccatc 900  
OY 3855 cgtatggcgcgaggtcgaacactggtcccgagttggtcggtcgtcgtcgtcgtcgtc 3914  
DB 901 cgtatggcgcgaggtcgaacactggtcccgagttggtcggtcgtcgtcgtcgtcgtc 960  
OY 3915 ggtatgaatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3974  
DB 961 ggtatgaatctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1020  
OY 3975 atgagcgcgcaaaactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4034  
DB 1021 atgagcgcgcaaaactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1080  
OY 4035 aagatgttcgttgagacagcttcaacagattgtcctcaaacctgcgtcgtcgtcgtc 4094  
DB 1081 aagatgttcgttgagacagcttcaacagattgtcctcaaacctgcgtcgtcgtcgtc 1140  
OY 4095 aaccattatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4154  
DB 1141 aaccattatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1200  
OY 4155 agtctgactaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4214  
DB 1201 agtctgactaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260  
OY 4215 ggtccaaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4274  
DB 1261 ggtccaaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320  
OY 4275 ttcccaagcgcgagtgatcgaactcgaactcgaactcgaactcgaactcgaactcgaact 4334  
DB 1321 ttcccaagcgcgagtgatcgaactcgaactcgaactcgaactcgaactcgaactcgaact 1380  
OY 4335 ggcagcaaaaggagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4394  
DB 1381 ggcagcaaaaggagatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1440  
OY 4395 ggcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4454  
DB 1441 ggcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1500  
OY 4455 gggaaagtttgagtttgaagttatcccgcaataggttaaatgtcgtcgtcgtcgtc 4514  
DB 1501 gggaaagtttgagtttgaagttatcccgcaataggttaaatgtcgtcgtcgtcgtc 1560  
OY 4515 atgtgtaataatggttgagttatgaataacacgtcgtcgtcgtcgtcgtcgtcgtc 4574  
DB 1561 atgtgtaataatggttgagttatgaataacacgtcgtcgtcgtcgtcgtcgtcgtc 1620  
OY 4575 gtgtgacacgcaacatcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 4634  
DB 1621 gtgtgacacgcaacatcactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1680



OY	4635	atgaaactctggggcgcattgggtttatagatccttcgcagaatgacagagtggttccctccattt	4694		
Db	1681	atgaatctgggggcattgggtttatagatccttcgcagaatgacagagtggttccctccattt	1740		
OY	4695	caagaataataccgaagaaccttaacacgaagatataggttaagtgatgaataactg	4754		
Db	1741	caagaataataccgaagaaccttaacacgaagatataggttaagtgatgaataactg	1800		
OY	4755	aaatgaacctcgaatactcgcgcattagagcgttgaagttcggacttaacgcattcatt	4814		
Db	1801	aaatgaacctcgaatactcgcgcattagagcgttgaagttcggacttaacgcattcatt	1860		
OY	4815	ctgatacgcaggagagatcgcgcacaaatgycgcgcatactgcccgaagcgtcttcaattcat	4874		
Db	1861	ctgatacgcaggagagatcgcgcacaaatgycgcgcatactgcccgaagcgtcttcaattcat	1920		
OY	4875	caccgatgagagagagaacacgcgcgcctcttaaaaaaagaatggtggaagatattctgat	4934		
Db	1921	caccgatgagagagagaacacgcgcgcctcttaaaaaaagaatggtggaagatattctgat	1980		
OY	4935	ccctggcttcaggagcctctgcctttaaaacctgaa	4969		
Db	1981	ccctggcttcaggagcctctgcctttaaaacctgaa	2015		
RESULT 5					
ID	AA518308	AA518308 standard; DNA: 5859 BP.			
AC	AA518308;	\			
XX					
DT	12-MAR-2002	(first entry)			
XX					
DE	DNA sequence of shuttle vector plasmid pADM291-4.				
XX					
KW	Cloning vector; Ketogulonigenium replicon; endogenous plasmid; mutant;				
KM	transformed host cell; Escherichia coli; pADM291-4; circular; cyclic;				
KX	shuttle vector plasmid; ds.				
OS	Ketogulonigenium sp. strain ADM291-19.				
OS	Synthetic.				
PN	MO20017347-A2.				
XX					
PD	18-OCT-2001.				
XX					
PF	05-APR-2001; 2001WO-US11059.				
XX					
PR	05-APR-2000; 2000US-194625P.				
XX					
PA	(ARCH ) ARCHER-DANIELS MIDLAND CO.				
XX					
PI	(DELT/) D'ELIA J.				
XX					
DR	D'Elia J;				
XX					
XX	WPI, 2002-049150/06.				
XX					
PT	Novel nucleic acid vector comprising Ketogulonigenium replicon found on				
PT	a specific deposited endogenous plasmid, useful for producing				
PT	polypeptides and/or transcripts by culturing host cells transformed				
XX					
PS	Claim 3; Fig 3; 66pp; English.				
XX					
CC	The present invention relates to the isolation of vectors comprising				
CC	a Ketogulonigenium replicon found on the endogenous plasmid, pADM291.				
CC	The invention also describes methods of transforming host cells with				
CC	the vectors and producing polypeptides and/or antisense transcripts by				
CC	culturing the transformed host cells. The vectors are useful for				
CC	transforming a host cell by conjugation or electroporation.				
CC	The vectors which have a replicon functional in both Ketogulonigenium				
CC	and Escherichia coli, enable the cloning of certain genes of				
CC	Ketogulonigenium in E.coli as the latter is an efficient host for				

CC amplification of vector DNA. The present DNA sequence represents the  
CC shuttle vector pAD291-4.  
XX  
SQ Sequence 5859 BP; 1444 A; 1487 C; 1566 G; 1362 T; 0 other;

Query Match 23.6%; Score 2007; DB 24; Length 5859

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2955 ggcgaatgggtcgaaattcatatgaattttgtgtgaggtgcgtacgcgcctctgacaggggtg 3014

```

Db 402 ggcgaatggctcgaatcatagaatttgtgtcaggtcgtagcgctctgacagggtg 461
      |||

```

QY 3015 ctgcgcggaatctctgtctcaggtaggcgcaaatgagaggtgttagttgccccctg 3074

Db 462 ctgcgcgaatctctgtctcagtaggcgcgaatgtagagtgcttagttgccccctg 521

QY 3075 tatcgcctctctgcgtgagcattggtcatccctgcccgcgacatatgatatattccgcctagag 3134

```
Db      522 tatcgtctctgctgtagcatgggatcatccgcgccgagatatgatattccgcctagg 581
```

**OY**    3135 gattactgataagttctcgcgtgtcgacctgttcggaccttgcgcgccttgatccgacattgtc 3194

Db 582 gattactgataagttctctgcctgtcggcctgttcgcggcctgttcgcggcctgttc 641

QY 3195 ggagcctgtccctcttgtccgcgctgtctcactttttcacatcaaaaatgggcgaagc 3254

Db 642 ggagcctgtccctctgtccgcctgtccctacttttcaacatcaaaaaatggcggaagc 701

QY 325 cctcttggtctatagttccttatagttacacgaaaattacacataattatcaatagctt 3114

Db 702 cctctctgtctatagttccttatagttcatacgaataattacacataattatcaatgctt 761

QY 3315 atcgcctaaaggagtaatttggccgcgcaaaaggagtaatttggccgcgcaaaaggagt 3374

Db 762 atcgcctaaaggagtaattggccgcgaaggagtaattggccgcgaaggagat 821

QY 3375 aatggcgccgcaaaagugagtaattggcgccgatactcggttggttacatggggaggaatcc 3434

Db 822 aattggccgcaaaagagtaattcggccgacatcggtctgttacacgggaggaatcc 881

QY 3435 ccttaatcatcttccccaatgggaagacacacacaagtggccgcagaccgcgccttcgac 3494

Db 882 cctaatcatcttcccatggaagacacacaaagtgcgcgcagaccggccttcgac 941

QY 3495 cagacaaactgtgtcctccctgcgaggttgcgagaggggtctatatgtcccatccgcc 3554

Db 942 cagacaaactgtgtccctgcgaggttggcgagaggggtatatatgscatccgcc 1001

3555 cgccctgcagcgctcaagctcatgcatcttaatgatagccactgcggcgcgcacatgct 3614

Db 1002 cgctgcagcgctcaagctcatgcatltaatgatagccacgcggcggccgcatgct 1061

3615 gatgatgtgcgccatgaaatgcggtcgcgcgacattcgcgcgaatcgacgcgcatgaaaac 3674

Db 1062 gatgatgtgcccataatgcgcgcgcgcgaattcgcgcgaatcgacgcgcattgaaaaaac 112121

QY 3675 catgacgctgagagcctgacccgcgtgtcgagagctagccgctgcggtgttgaccat 3734

Db 1122 catgacgctgagagccctgaccccgctgcttgagagagctagccgctgscgtgttgaccat 1181

3735 gatgacccctgcaagatgatcgtgcacgtcgcgcgttctgtcgtatgagcgcgaaatagac 37394

Db 1182 gatgacccctgaagatgatcgtgacagtcgcgcttgctcgtatgagcgcgaaatagac 1241

QY 3795 taccgcagcagcgcaagcggcgaactcctagtgacgtgacgttcgcgagtlacattcgt 3854

Db 1242 taccgccaagcgcaactcctagtacgtgaccttcgaglacatccgt 1301

Oy 3855 cgtatggcgcgcgagtcgaaccactgggcattctcgaccgtcaaacggtattccatctc 3914

Db 1302 cgatctgcgcgcgagtcgaaccactcgggcattctctgcacgcgtcaaacggtattccatctc 1361



QY 3915 ggtatgaatattcgtgctgctgttccagcagctctagtctcgcaatctgacg 3974  
 |||||  
 Db 1362 ggtatgaatattcgtgctgctgttccagcagctctagtctcgcaatctgacg 1421  
 QY 3975 atgagcgcaaaaccttaccggtcccgaggtcgcggtctcttgaagtcccgaggga 4034  
 |||||  
 Db 1422 atgagcgcaaaaccttaccggtcccgaggtcgcggtctcttgaagtcccgaggga 1481  
 QY 4035 aagatggttcgttggaaagcagcttaacagatttgccttaaaccttcagatagatc 4094  
 |||||  
 Db 1482 aagatggttcgttggaaagcagcttaacagatttgccttaaaccttcagatagatc 1541  
 QY 4095 aacattatcgcgtctgcagattgcagcgcaagcgcaacagatttgccttgcgttgcga 4154  
 |||||  
 Db 1542 aacattatcgcgtctgcagattgcagcgcaagcgcaacagatttgccttgcgttgcga 1601  
 QY 4155 agtctgactataagcttggaaagtgaagacgaccccaacgctcgcaaggcgagctggcg 4214  
 |||||  
 Db 1602 agtctgactataagcttggaaagtgaagacgaccccaacgctcgcaaggcgagctggcg 1661  
 QY 4215 gttcccaagtcggttcgagatgctcgttcgcaagagcgacgggaaacgataagccccctcc 4274  
 |||||  
 Db 1662 gttcccaagtcggttcgagatgctcgttcgcaagagcgacgggaaacgataagccccctcc 1721  
 QY 4275 ttcccaagcgagcgaggtacactacagtcacgctgtgtgttggagctgaagcgtctgct 4334  
 |||||  
 Db 1722 ttcccaagcgagcgaggtacactacagtcacgctgtgtgttggagctgaagcgtctgct 1781  
 QY 4335 ggcacaaacagaggaacagatctgacgtccagactccggttcttctcgaggagaga 4394  
 |||||  
 Db 1782 ggcacaaacagaggaacagatctgacgtccagactccggttcttctcgaggagaga 1841  
 QY 4395 ggcgtgctcgttcgagcgttcgcaaacatcgaaaaactgttttaagatttcttgcgcaaggtta 4454  
 |||||  
 Db 1842 ggcgtgctcgttcgagcgttcgcaaacatcgaaaaactgttttaagatttcttgcgcaaggtta 1901  
 QY 4455 gggaaaggttcttgaagtttccagatctccagatccggttcttctcgaggagaga 4514  
 |||||  
 Db 1902 gggaaaggttcttgaagtttccagatctccagatccggttcttctcgaggagaga 1961  
 QY 4515 atgtgaatatagcgttgaagatatagaatatacagcgttcgagacggtcgcaaaagcagcg 4574  
 |||||  
 Db 1962 atgtgaatatagcgttgaagatatagaatatacagcgttcgagacggtcgcaaaagcagcg 2021  
 QY 4575 gtgtgctgacgcgcaacatcattcgtcggtcggttaaaagcggttaaaattcccgtaaaaag 4634  
 |||||  
 Db 2022 gtgtgctgacgcgcaacatcattcgtcggtcggttaaaagcggttaaaattcccgtaaaaag 2081  
 QY 4635 atgaatctgggcatggttatagatccttcgcaagatttgcagagagtttcttccattc 4694  
 |||||  
 Db 2082 atgaatctgggcatggttatagatccttcgcaagatttgcagagagtttcttccattc 2141  
 QY 4695 caaagaaatacacgcgaaaccccttaacagcgaagtatatgttaagcgttgaatgaacacatg 4754  
 |||||  
 Db 2142 caaagaaatacacgcgaaaccccttaacagcgaagtatatgttaagcgttgaatgaacacatg 2201  
 QY 4755 aaatgaactcaagaatacagcagcatatagagcgttgaagttcggacttcaagcgtattat 4814  
 |||||  
 Db 2202 aaatgaactcaagaatacagcagcatatagagcgttgaagttcggacttcaagcgtattat 2261  
 QY 4815 ctgattccagaggagatcgacgacaatagcgcgacatggcgagcgtcttcaattcat 4874  
 |||||  
 Db 2262 ctgattccagaggagatcgacgacaatagcgcgacatggcgagcgtcttcaattcat 2321  
 QY 4875 caccgatagaagaggaagacccgccccctcaaaaacaaagattggtgaagatatctgat 4934  
 |||||  
 Db 2322 caccgatagaagaggaagacccgccccctcaaaaacaaagattggtgaagatatctgat 2381  
 QY 4935 cctgggttcagagccttgcctttaaactga 4969  
 |||||  
 Db 2382 cctgggttcagagccttgcctttaaactga 2416

RESULT 6  
 AAV30458 standard; DNA: 534720 BP.  
 ID AAV30458 standard; DNA: 534720 BP.  
 XX  
 AC AAV30458;  
 XX  
 DT 14-Oct-1998 (first entry)  
 XX  
 DE Rhizobium species plasmid pNGR234a.  
 XX  
 KW Symbiosis; open reading frame; ORF; plasmid; vector; transportation;  
 degradation; metabolism; host range; nitrogen fixation; nodulation;  
 XX  
 KW Legume; plant; ds.  
 XX  
 OS Rhizobium sp.  
 XX  
 Key  
 FH  
 FT CDS  
 Location/Qualifiers  
 417796..418671  
 /tag= a  
 /standard\_name= "ORF K1"  
 /product= "oligopeptide permease"  
 /note= "homologous to the OppC gene"  
 418673..419680  
 /tag= b  
 /standard\_name= "ORF K2"  
 /product= "oligopeptide permease"  
 /note= "homologous to the OppD gene"  
 419677..420738  
 /tag= c  
 /standard\_name= "ORF K3"  
 /product= "oligopeptide permease"  
 /note= "homologous to the OppF gene"  
 420774..422159  
 /tag= d  
 /standard\_name= "ORF K4"  
 /product= "encapsulation-like protein"  
 /note= "homologous to the CapA gene"  
 422628..424031  
 /tag= e  
 /standard\_name= "ORF K5"  
 /product= "aminotransferase-like protein"  
 /note= "homologous to the BioA gene"  
 424056..425594  
 /tag= f  
 /standard\_name= "ORF K6"  
 /product= "(semi)aldehyde dehydrogenase-like protein"  
 complement (426949..428028)  
 /tag= g  
 /standard\_name= "ORF K7"  
 /product= "transposase homologue"  
 /note= "homologous to the Tnp gene"  
 428292..429623  
 /tag= h  
 /standard\_name= "ORF K8"  
 /product= "glutamate dehydrogenase-like protein"  
 /note= "homologous to the GluDI gene"  
 complement (430538..431284)  
 /tag= i  
 /standard\_name= "ORF K9"  
 /product= "transposase homologue"  
 complement (431296..432840)  
 /tag= j  
 /standard\_name= "ORF K10"  
 /product= "transposase homologue"  
 /note= "homologous to the Tnp gene"  
 complement (433880..434110)  
 /tag= k  
 /standard\_name= "ORF K11"  
 /product= "protein of unknown function"  
 /note= "homologous to the Fixu gene"  
 complement (434107..434433)  
 /tag= l  
 CDS



```

FT      /standard_name="ORF K12"
FT      /product="protein of unknown function"
FT      complement (434517..434711)
FT      /tag= m
FT      /standard_name="ORF K13"
FT      /product="ferrodoxin/ferrodoxin-like protein"
FT      /note="homologous to the fdxn gene"
FT      complement (434753..436234)
FT      /tag= n
FT      /standard_name="ORF K14"
FT      /gene="nifB"
FT      /product="protein involved in FxMo co-factor
FT      biosynthesis"
FT      complement (436460..438130)
FT      /tag= o
FT      /standard_name="ORF K15"
FT      /gene="nifA"
FT      /product="positive regulator of nif, fix and other
FT      genes"
FT      complement (438297..438590)
FT      /tag= p
FT      /standard_name="ORF K16"
FT      /gene="fixX"
FT      /product="protein required for nitrogenase activity"
FT      complement (438605..439912)
FT      /tag= q
FT      /standard_name="ORF K17"
FT      /gene="fixC"
FT      /product="protein required for nitrogenase activity"
FT      complement (439923..441032)
FT      /tag= r
FT      /standard_name="ORF K18"
FT      /gene="fixB"
FT      /product="protein required for nitrogenase activity"
FT      complement (441042..441899)
FT      /tag= s
FT      /standard_name="ORF K19"
FT      /gene="fixA"
FT      /product="protein required for nitrogenase activity"
FT      complement (442316..442636)
FT      /tag= t
FT      /standard_name="ORF K20"
FT      /product="protein of unknown function"
FT      complement (443313..443879)
FT      /tag= u
FT      /standard_name="ORF K21"
FT      /product="protein of unknown function"
FT      444337..445029
FT      /tag= v
FT      /standard_name="ORF K22"
FT      /product="ferrodoxin-like protein"
FT      /note="homologous to the NifQ gene"
FT      445088..446602
FT      /tag= w
FT      /standard_name="ORF K23"
FT      /gene="dctA"
FT      /product="C4-dicarboxylate transport protein"
FT      /note="homologous to the DctAI gene"
FT      446599..447843
FT      /tag= x
FT      /standard_name="ORF L1"
FT      /product="cytochrome P450-like protein"
FT      /note="homologous to the Camc gene"
FT      447844..448500
FT      /tag= y
FT      /standard_name="ORF L2"
FT      /product="gamma-hexachlorocyclohexane-dechlorinase-like
FT      protein"
FT      448497..450203
FT      /tag= z
FT      /standard_name="ORF L3"
FT      /product="putative protein with degradative function"

```

```

FT      CDS
FT      450341..451396
FT      /tag= aa
FT      /standard_name="ORF L4"
FT      /product="luciferase alpha-subunit-like protein"
FT      /note="homologous to the LuxA gene"
FT      452980..454494
FT      /tag= ab
FT      /standard_name="ORF L6"
FT      /gene="nifD"
FT      /product="alpha-subunit of FxMo protein of nitrogenase"
FT      454590..456131
FT      /tag= ac
FT      /standard_name="ORF L7"
FT      /gene="nifX"
FT      /product="beta-subunit of FxMo protein of nitrogenase"
FT      456187..457677
FT      /tag= ad
FT      /standard_name="ORF L8"
FT      /product="protein involved in FxMo co-factor
FT      biosynthesis"
FT      /tag= ae
FT      /standard_name="ORF L9"
FT      /product="protein involved in FxMo co-factor
FT      biosynthesis"
FT      /note="homologous to the NifB gene"
FT      457687..459096
FT      /tag= ae
FT      CDS
FT      /standard_name="ORF L9"
FT      /product="protein involved in FxMo co-factor
FT      biosynthesis"
FT      /note="homologous to the fixF gene"
FT      459093..459575
FT      /tag= af
FT      /standard_name="ORF L10"
FT      /product="protein of unknown function"
FT      /note="homologous to the NifX gene"
FT      459579..460067
FT      /tag= ag
FT      /standard_name="ORF L11"
FT      /product="protein of unknown function"
FT      460501..460920
FT      /tag= ah
FT      /standard_name="ORF L12"
FT      /product="protein similar to part of the Fe protein
FT      of nitrogenase"
FT      /note="homologous to the NifH gene"
FT      461228..461545
FT      /tag= ai
FT      CDS
FT      /standard_name="ORF L13"
FT      /product="protein of unknown function"
FT      463201..464739
FT      /tag= aj
FT      /standard_name="ORF L14"
FT      /product="peptidase-like protein"
FT      /note="homologous to the bt-MPP gene"
FT      464736..466079
FT      /tag= ak
FT      /standard_name="ORF L15"
FT      /product="processing protease-like protein"
FT      /note="homologous to the pp gene"
FT      466590..467021
FT      CDS

```

Query Match 0.8%; Score 65; DB 19; Length 534720;  
Best Local Similarity 55.6%; Pred. No. 1.9e-05;  
Matches 125; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

Oy      1809 aaccatgagagccatctcgagagcgatcccgccgaagaacagacggttctacc 1868
Oy      1869 gcatgacgtctcgcaacgcttcggggaatgagacgtatcgagatgggcgca 1928
Db      83624 ccatgtgatcgaaaccccttcgcgcgaacatctgctgttcgcaagtgggctgca 83663
Oy      1929 ttgagcatcagcgccgcatccgctcgaltgtgttgagatgaaacaagatgccatcgctg 1988
Db      83684 ttggaaccaagagacatgatgattccatcatttcggccaagaagaagatggctccgat 83743

```



Db 83684 ttgga

Method	Best Local Similarity
Exact	49.14

•























XX	
PN	WO200159126-A2.
XX	
PD	16-AUG-2001.
XX	
PF	08-FEB-2001; 2001WO-GH0509.
XX	
PR	08-FEB-2000; 2000GB-0002840.
XX	
PR	10-APR-2000; 2000GB-0008786.
XX	
PR	14-APR-2000; 2000GB-0009387.
XX	
PX	(UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA	(SNTF) SINTEF STIPELSEN IND TEK FORSK.
PA	(ALPH-) ALPHARMA AS.
PA	(SINV-) SINVENT AS.
PA	(DZIE/) DZIELEWSKA H.
PA	(ZOTC/) ZOTCHEV S B.
PA	(SEKU/) SEKUROVA O N.
PA	(FUAE/) FJAEVRVI E.
PA	(BRAU/) BRAUTASET T.
PA	(STRO/) STROM A R.
XX	
XZ	Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
PI	Valla S, Ellingsen TE, Sletta H, Guilleksen O;
XX	
DR	WPI: 2001-557614/62.
XX	
DR	P-PDSB: AAE10143, AAE10144, AAE10145, AAE10146, AAE10147, AAE10148,
XX	AAE10149, AAE10150.
PT	New nystatin polyketide synthase polynucleotides and polypeptides,
XX	useful as antibiotics and antifungals -
PS	Claim 1; Page 188-254; 266pp; English.
CC	The present invention relates to the cloning and sequencing of the gene
CC	cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC	involved in the biosynthesis of the macrolide antibiotic nystatin.
CC	The nystatin PKS is useful as antifungal antibiotics. The present
CC	sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.
XX	
SQ	Sequence 125401 BP; 15664 A; 49632 C; 42871 G; 17174 T; 0 other;

[illegible]

Db 39012 cggccgcgctctcgtccccgcgcgcgcgaacccccctggcgccctgggcagcgcgggcacaagg 39071

QY 1633 cagcctctacggcgtcgaactgcgtcgcgcgaacctccccagcc 1676

Db 39072 cagcctcgacgcgcctgcgcctctcgtcccaaccgcgcgagcgccgac 39115

Search completed: August 20, 2002, 13:15:54  
Job time: 20837 sec



**THIS PAGE BLANK (USPTO)**

---



**THIS PAGE BLANK (USPTO)**



QY	1689	agccgcgaatctgcygcgcgcacacactgtctgtgttgccatgacagatctgagagacaca	1748
Db	2085	agacgcgtcccggtgtgtctcccgctcgcgtccgcgtcccccacccgcctccatcacgaca	2144
QY	1749	aggaatatactgcgacccgagcaagccgagggccctcgatatcgacagagatacactct	1808
Db	2145	aggaagagacaagacgacccatctgcgtctcctgtgacccctgcgttcgaatcacacgacact	2204
QY	1809	aaccatctggaacctactctgagaaagcgacatcccgccgaagacaacagacggtctaac	1868
Db	2205	acgcgcctgtgtccccctgctctgtggtcgctgtcgccgcgcgcagaggtgtgcgtctgcaca	2264
QY	1869	gcattgycggtccctgcgcgaacctgtctggggaatgtgacgcgtctatcgagaatggtgcga	1928
Db	2265	gccaaaccgcgcctacacgagacaacataccgcggtgcgcggtctgcgcgcgtgtccgtctgcga	2324
QY	1929	tctgagatcaagcgcgacatccgcgtctgattgtgtctgtagagtgaaacaagatgcacgtgtg	1988
Db	2325	ccgacacactcatctcacagagttacgcggtgtggtgagtgtcggtgcgcgcgcgcgcgaacatc	2384
QY	1989	cgatgctctgcacatctgaaacccgcgaagcgtctagcgcggtctattgtgctctgagcccatccaga	2048
Db	2385	cggcgatgcctcttcgaacgagagccgcgtctccgaagccgcgtcttgactctgtggaacacgcctcttcgaca	2444











```

CORRESPONDENCE ADDRESS:
ADDRESS: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,092
FILING DATE: 20-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/047,221
FILING DATE: 20-MAY-1997
APPLICATION NUMBER: 60/060,465
FILING DATE: 30-SEP-1997
APPLICATION NUMBER: 60/075,940
FILING DATE: 25-FEB-1998
APPLICATION NUMBER: 60/077,033
FILING DATE: 06-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7032
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 733 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-082-092-15

```

```

Query Match          0.5%; Score 41.4; DB 4; Length 733;
Best Local Similarity 47.5%; Pred. No. 0.15;
Matches 123; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 1468 ggaagggacacgcgagagcgcggtctatctgcgcgcctctgacaaagccgcgtgcacag 1527
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 467 ggagggaaggggggcgagggcgagcccttgccggaggcgccgagcccttgagcatttcg 526

QY 1528 ctgagctgtcacccgagagggcgcctgcgtgtccacacccaccttaagcgaagacagct 1587
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 527 ctgcctgattccggcgcacagagattttccccagctgcgcgcacacagcagccggcg 586

QY 1588 gtgagcctgttcaagctcgtcgtcgaagaagcgcgctcgtatccagcctctacggct 1647
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 587 gccctgcctcgggccccgctgcgcgcgcgcgagagagctgctgagagcgcagccagccca 646

QY 1648 gaactcgtgcgcgaaccttccacagcaccatctacgcgcgagacgcggaatctbgcgcg 1707
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 647 gccggcgccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 706

QY 1708 cgcacagctcgtcgtgcgc 1726
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 707 ccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 725

```

```

RESULT 7
US-08-295-060-3
Sequence 3, Application US/08295060
Patent No. 5659123
GENERAL INFORMATION:
APPLICANT: VAN RIE, Jeroen

```

```

APPLICANT: JANSSENS, Stefan
APPLICANT: PEREROEN, Marix
TITLE OF INVENTION: NEW DIABROTICA TOXINS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,060
FILING DATE: 26-AUG-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1957 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: YES
AMT1-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1947
US-08-295-060-3

```

```

Query Match          0.5%; Score 40.8; DB 1; Length 1957;
Best Local Similarity 50.5%; Pred. No. 0.41;
Matches 99; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1482 agagccctgcggtctatctgcgcgcctctgacaaagccgcgtgcacagctgctgttcacg 1541
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 416 agggcctgcagaaacagctgagagactgagcgcgccttgagcagctgcagagaagacc 475

QY 1542 gacagggcctcgcgtgttcccaacccaccttaagcgaagacagcgtgtgagcgttca 1601
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 476 ccgtgacacacacccacaccccccagcagcagcagcagcagcagcagcagcagcagc 535

QY 1602 agtctcgtcgcgaagaagcgcggtgctgctccagcctctacggcgctgcgcgcgcgcgc 1661
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 536 agaaccacttccgacaaacagcagcagcagcagcagcagcagcagcagcagcagcagc 595

QY 1662 gaaccttcccaagcca 1677
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 596 tgaccacttaccgcca 611

```

```

RESULT 8
US-08-464-517-5/c
Sequence 5, Application US/08464517
Patent No. 5869640
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD

```



```

: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: . US/08/464,517
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/963,308
: FILING DATE: 16-OCT-1992
: APPLICATION NUMBER: US 07/888,178
: FILING DATE: 26-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/701,514
: FILING DATE: 16-MAY-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Matthew P. Vincent
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: MII-004C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1926 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 101..940
:
: US-08-464-517-5

```

```

Query Match          0.5%; Score 40.4; DB 2; Length 1926;
Best Local Similarity 47.6%; Pred. No. 0.53;
Matches 119; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

```

```

QY 1549 catcgctggtccacaccacacttagcgagagccagctgtgscgtgtcaagtcctg 1608
   || || || || || || || || || || || || || || || || || || || ||
DB 295 CAGCATCCAGTAGGACGACGATCTTCCGCATGTGCGGCTTGATCTCCCTGCACGCACTG 236
   || || || || || || || || || || || || || || || || || || || ||
QY 1609 gctcgagaagcgcggtcgtccagcctctacgggtgtgacgtgctgscgcaactt 1668
   || || || || || || || || || || || || || || || || || || || ||
DB 235 GAAGTAGAGGCGCGGGGTAGTAGCGCTCTCCACAGCGGAGCAGGCTCTGCAGACACG 176
   || || || || || || || || || || || || || || || || || || || ||
QY 1669 ccccgacacatctaccgcgagaccggaatctgscgcgcacacagctgtctg9gcca 1728
   || || || || || || || || || || || || || || || || || || || ||
DB 175 CTGCTCCCCCAGACGCGCGGGGTCTCGGCCGCGGCGCGGTCCCGGTGCTTCGCA 116
   || || || || || || || || || || || || || || || || || || || ||
QY 1729 tgcagcatcgagagcaccgaagagtagatcgagcgcgagcaagcgagccctgatat 1788
   || || || || || || || || || || || || || || || || || || || ||
DB 115 ACACAGCAGCTCCATCTCTGGGCGAGCAAGCGAGGCGGAGTGGCGCTTCGCGAGTTC 56
   || || || || || || || || || || || || || || || || || || || ||
QY 1789 cgacagagag 1798
   || || || || || || || || || || || || || || || || || || || ||
DB 55 CCAAGGCAGG 46

```

```

RESULT 9
US-08-463-772-5/c
: Sequence 5, Application US/08463772
: Patent No. 6066501
: GENERAL INFORMATION:
: APPLICANT: BEACH, David H.

```

```

: TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
: NUMBER OF SEQUENCES: 50
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/463,772
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/963,308
: FILING DATE: 16-OCT-1992
: APPLICATION NUMBER: US 07/888,178
: FILING DATE: 26-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/701,514
: FILING DATE: 16-MAY-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Matthew P. Vincent
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: MII-004C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1926 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 101..940
:
: US-08-463-772-5

```

```

Query Match          0.5%; Score 40.4; DB 3; Length 1926;
Best Local Similarity 47.6%; Pred. No. 0.53;
Matches 119; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

```

```

QY 1549 catcgctggtccacaccacacttagcgagagccagctgtgscgtgtcaagtcctg 1608
   || || || || || || || || || || || || || || || || || || || ||
DB 295 CAGCATCCAGTAGGACGACGATCTTCCGCATGTGCGGCTTGATCTCCCTGCACGCACTG 236
   || || || || || || || || || || || || || || || || || || || ||
QY 1609 gctcgagaagcgcggtcgtccagcctctacgggtgtgacgtgctgscgcaactt 1668
   || || || || || || || || || || || || || || || || || || || ||
DB 235 GAAGTAGAGGCGCGGGGTAGTAGCGCTCTCCACAGCGGAGCAGGCTCTGCAGACACG 176
   || || || || || || || || || || || || || || || || || || || ||
QY 1669 ccccgacacatctaccgcgagaccggaatctgscgcgcacacagctgtctg9gcca 1728
   || || || || || || || || || || || || || || || || || || || ||
DB 175 CTGCTCCCCCAGACGCGCGGGGTCTCGGCCGCGGCGCGGTCCCGGTGCTTCGCA 116
   || || || || || || || || || || || || || || || || || || || ||
QY 1729 tgcagcatcgagagcaccgaagagtagatcgagcgcgagcaagcgagccctgatat 1788
   || || || || || || || || || || || || || || || || || || || ||
DB 115 ACACAGCAGCTCCATCTCTGGGCGAGCAAGCGAGGCGGAGTGGCGCTTCGCGAGTTC 56
   || || || || || || || || || || || || || || || || || || || ||
QY 1789 cgacagagag 1798
   || || || || || || || || || || || || || || || || || || || ||
DB 55 CCAAGGCAGG 46

```

```

RESULT 10
PCT-US93-05000-5/c

```



```

; Sequence 5, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSH91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1926 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-05000-5

```

```

Query Match 0.5%; Score 40.4; DB 5; Length 1926;
Best Local Similarity 47.6%; Pred. No. 0.53;
Matches 119; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1549 catcgcgtgtccacacccacctagcgaagcagctgtgagcctgttcaagtcctg 1608
DB 295 CAGCATCCAGTAAGCCACATCTTCCGATGTGCGCTTGATCTCCGCTGCAGCAGCTG 236
QY 1609 gctcgaagaagcgcgcgtcgaatccagcctcagcgcgtcgaactcgtcgcgaacct 1668
DB 235 GAAGTAGAGGCGGGGTAGTAGCGCTCTCCAGGGGAGCAGGCTCTGCAGAGACG 176
QY 1669 ccccaagcacatctacgcgcgaagcgaactcgcgcgcgaagctgtcgtcgcgaacct 1728
DB 175 CTGTGTCCTCCAGCAGCGCGGCTCCGCCCGCGGCGCGGCGGCTGCTTCCCA 116
QY 1729 tgcgaagcagcagcagcgaagagatatacgcgcgaagcgaagcgcgcgcctcgatat 1788
DB 115 ACACAGCAGCTCCATCTCTGGGACGACGACAGCAGGCGGAGTGCCTGCGGAGTC 56
QY 1789 cgcacgagag 1798
DB 55 CCAAGGCAGG 46

RESULT 11
US-08-246-361A-5/c
; Sequence 5, Application US/08246361A
; Patent No. 5998582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

```

```

; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIT-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 101..976
; US-08-246-361A-5

```

```

Query Match 0.5%; Score 40.4; DB 2; Length 1962;
Best Local Similarity 47.6%; Pred. No. 0.54;
Matches 119; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1549 catcgcgtgtccacacccacctagcgaagcagcgtgtgagcctgttcaagtcctg 1608
DB 295 CAGCATCCAGTAAGCCACATCTTCCGATGTGCGCTTGATCTCCGCTGCAGCAGCTG 236
QY 1609 gctcgaagaagcgcgcgtcgaatccagcctcagcgcgtcgaactcgtcgcgaacct 1668
DB 235 GAAGTAGAGGCGGGGTAGTAGCGCTCTCCAGGGGAGCAGGCTCTGCAGAGACG 176
QY 1669 ccccaagcacatctacgcgcgaagcgaactcgcgcgcgaagctgtcgtcgcgaacct 1728
DB 175 CTGTGTCCTCCAGCAGCGCGGCTCCGCCCGCGGCGCGGCGGCTGCTTCCCA 116
QY 1729 tgcgaagcagcagcgaagagatatacgcgcgaagcgaagcgcgcgcctcgatat 1788
DB 115 ACACAGCAGCTCCATCTCTGGGACGACGACAGCAGGCGGAGTGCCTGCGGAGTC 56
QY 1789 cgcacgagag 1798
DB 55 CCAAGGCAGG 46

```

```

RESULT 12
US-08-457-245-4/c

```



```

: Sequence 4, Application US/08457245
: Patent No. 5573915
: GENERAL INFORMATION:
: APPLICANT: BARRY III, Clifton E.
: APPLICANT: YUAN, Ying
: TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED
: TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN
: TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
: STREET: Stewart Street Tower, One Market Plaza
: CITY: San Francisco
: STATE: California
: COUNTRY: US
: ZIP: 94105-1493
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/457,245
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Chambers, Guy W.
: REGISTRATION NUMBER: 30,617
: REFERENCE/DOCKET NUMBER: 15280-216000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 543-9600
: TELEFAX: (415) 543-5043
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1155 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: complement (4..1155)
: OTHER INFORMATION: /product= "ORF2"
:
: US-08-457-245-4

```

```

Query Match          0.5%: Score 39.6; DB 1; Length 1155;
Best Local Similarity 45.7%; Pred. No. 0.64;
Matches 138; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1183 gacgacacccacccaagcgctggaacaaagcgcggtgtgctgggaaanaaacgcgcgct 1242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 637 gtcgattacgcgcgtcggtggtgacacacacacacacacacacacacacacacacacacac 578

QY 1243 gacacccgacacagattgcctgcatctcattcctgcgcgaagaaagcggtgtgctgggaa 1302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 577 ggcggtacacttacttgcgcggcggtggtggtggtggtggtggtggtggtggtggtggtg 518

QY 1303 tctggtcctgtcaagctgacagcagacacagcttgcgcgcgcgcgcgcgcgcgcgcgcgc 1362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 517 ggcgcgcgttggccacacacacacacacacacacacacacacacacacacacacacacacac 458

QY 1363 ggcgcgtcgtgattgctgacacacacacacacacacacacacacacacacacacacacacac 1422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 457 gtacagctcgtatctgcccacacacacacacacacacacacacacacacacacacacacacac 398

QY 1423 gaagacggagggccgcgaatctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 397 cgaacacccctgtcgacacacacacacacacacacacacacacacacacacacacacacacac 338

QY 1483 ga 1484

```

```

Db          337 GA 336

RESULT 13
US-08-457-245-1/c
: Sequence 1, Application US/08457245
: Patent No. 5573915
: GENERAL INFORMATION:
: APPLICANT: BARRY III, Clifton E.
: APPLICANT: YUAN, Ying
: TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED
: TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN
: TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
: STREET: Stewart Street Tower, One Market Plaza
: CITY: San Francisco
: STATE: California
: COUNTRY: US
: ZIP: 94105-1493
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/457,245
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Chambers, Guy W.
: REGISTRATION NUMBER: 30,617
: REFERENCE/DOCKET NUMBER: 15280-216000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 543-9600
: TELEFAX: (415) 543-5043
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5100 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 521..1381
: OTHER INFORMATION: /product= "cma"
:
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 1388..2539
: OTHER INFORMATION: /product= "ORF2"
:
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 4158..5027
: OTHER INFORMATION: /product= "ORF3"
:
: US-08-457-245-1

Query Match          0.5%: Score 39.6; DB 1; Length 5100;
Best Local Similarity 45.7%; Pred. No. 1.6;
Matches 138; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1183 gacacacccacccaagcgctggaacaaagcgcggtgtgctgggaaanaaacgcgcgct 1242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2021 gtcgattacgcgcgtcggtggtgacacacacacacacacacacacacacacacacacacac 1962

QY 1243 gacacctgacacagattgcctgcatctcattcctgcgcgaagaaagcggtgtgctgggaa 1302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1961 ggcggtacacttacttgcgcggcggtggtggtggtggtggtggtggtggtggtggtggtg 1902

```















GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 06:10:42 ; Search time 5789.92 Seconds  
(without alignments)  
19835.437 Million cell updates/sec

Title: US-09-826-206-2

Perfect score: 8509

Sequence: 1 tggtagacgcattgcttga.....tatgttgacgcgcgcgcga 8509

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	0.8	695	12	AQ289334 nbxb00341L
2	64.6	0.8	442	9	AM839280 CML-LT006
3	60.8	0.7	925	12	CNS0091P
4	54.4	0.6	925	12	CNS0091P
5	49.8	0.6	839	12	CNS004NB
6	48.6	0.6	421	9	AV701967
7	48.4	0.6	932	12	CNS00720
8	48	0.6	935	12	CNS006XK
9	47.2	0.6	518	9	AM671033
10	46.8	0.6	910	12	CNS0060N
11	46.4	0.5	494	10	BG384839
12	46.4	0.5	506	10	BG384302
13	46.4	0.5	512	10	BF078334
14	46.4	0.5	544	10	BG895635
15	46.4	0.5	558	9	AM784766
16	46.4	0.5	569	10	BI360602
17	46.4	0.5	1100	12	CNS016KD

18	46.2	0.5	428	12	BH386486
19	46.2	0.5	477	9	BE232046
20	46	0.5	206	10	BF962526
21	45.8	0.5	304	10	BF960900
22	45.6	0.5	1101	12	CNS01258
23	45.4	0.5	499	9	BE231844
24	45.2	0.5	441	10	BG417919
25	45.2	0.5	550	10	BG367643
26	45.2	0.5	580	9	BE194138
27	45.2	0.5	667	10	BE602055
28	45.2	0.5	710	10	BE455201
29	45.2	0.5	747	9	BE193612
30	45.2	0.5	812	10	BG369542
31	45.2	0.5	877	10	BG369006
32	45.2	0.5	883	10	BG417826
33	45.2	0.5	893	10	BG415257
34	45.2	0.5	924	10	BG365895
35	45.2	0.5	928	9	BE196050
36	45.2	0.5	1281	12	AG075401
37	45	0.5	650	10	BM488443
38	44.8	0.5	473	10	BE194372
39	44.8	0.5	1232	12	AG072425
40	44.6	0.5	230	10	C55473
41	44.4	0.5	269	10	BF947280
42	44.4	0.5	312	10	BF947371
43	44	0.5	392	9	AM497369
44	44	0.5	497	10	BJ204205
45	44	0.5	501	10	BJ183869

#### ALIGNMENTS

RESULT 1  
AQ289334 695 bp DNA linear GSS 03-DEC-1998  
nbxb00341L0r CUGI Rice BAC Library Oryza sativa genomic clone  
LOCUS nbxb00341L0r, DNA sequence.

ACCESSION AQ289334  
VERSION AQ289334.1  
KEYWORDS GI:3950780  
SOURCE GSS.

ORGANISM Oryza sativa.

REFERENCE Oryza sativa.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

AUTHORS 1 (bases 1 to 695)  
Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)

COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu

Seq primer: GGAAACACTATGACCAAC  
Class: BAC ends  
High quality sequence stop: 429.

#### FEATURES

source  
1..695  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbxb00341L0r"  
/clone\_lib="CUGI Rice BAC Library"  
/tissue\_type="leaf"  
/lab\_host="E. coli DH10B"  
/note="Vector: pBel0BAC11; Site\_1: HindIII; Site\_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those























	LOCUS	BG384839	494 bp	mRNA	linear	EST 12-MAR-2001
	DEFINITION	3041170 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.				
	ACCESSION	BG384839				
	VERSION	BG384839.1	GI:1309311			
	KEYWORDS	EST.				
	SOURCE	pig.				
	ORGANISM	Sus scrofa				
	AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus. 1.(bases 1 to 494) Fahrenkrug,S.C., Fieking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.				
	TITLE	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine				
	JOURNAL	Unpublished (2000)				
	COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -nummatch 12 options. PCR primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTPTCCAGTCACGACG Plate: 93 row: F column: 1 Seq primer: ATTAGGTGACATATGAT.  Location/Qualifiers 1..494 /organism="Sus scrofa" /db_xref="taxon:9823" /clone_lib="MARC 1Pig" /tissue_type="Pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."				
	BASE COUNT	90 a 174 c 155 g 75 t				
	ORIGIN					
	Query Match	0.5%;	Score 46.4;	DB 10;	Length 494;	
	Best Local Similarity	45.7%;	Pred. No. 0.64;			
	Matches 161;	Conservative 0;	Mismatches 191;	Indels 0;	Gaps 0;	
Oy	1458	ccgcgcttcggaggagggacacgagcagcgatcgcgtgtgccacaccaccttaagc	1517			
Dd	132	cctaacaagttttgaaacatgcctgatgacccctcgatgaactgaagtcggtagcggg	191			
Oy	1518	cgtctcacagctgtgctgttcacacgagacagggatcgcgtgtgccacaccaccttaagc	1577			
Dd	192	cagccacagaccctgctgtagaacccgagcgttgatgcgagccctggcgcatgaacctgacactact	251			
Oy	1578	agaagcagactgtgtgctgtgtccaagtlccttgtcctgagaaaggcgcgtgtgatccccagcc	1637			
Dd	252	tgtttccggagccggccacacatcaactcccattttgagaaaggcccgctgagcccgccgct	311			
Oy	1638	tctaaagcgtgacacgcgtgtgctgacacacttccacagcacatatcccgagacagcgca	1697			
Dd	312	tccggcgagagacacgcctgctgcccgcgtagccattccgggagagagcgttgcatcgcctgca	371			
Oy	1698	atctgtgcgcgcgcacagactgtctgtgtggccalcgcagatagagagaccaaaggaata	1757			
Dd	372	aacctttggagagccctgttcgacctgcccacagccatcacatcagagcccgagccgagccg	431			
Oy	1758	tggagacgagaaagcgcgagccctcgatatacagcagagatatacacttta	1809			
Dd	432	atggcagccgctgagacacgccttatgacatgacatgacacgaatgacatgca	483			

RESULT	12
LOCUS	BC384302 506 bp mRNA linear EST 12-MAR-2001
DEFINITION	303332 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION	BC384302
VERSION	BC384302.1 GI:13308774
KEYWORDS	EST.
SOURCE	pig.
ORGANISM	Sus scrofa
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus. 1 (bases 1 to 506) Fahrenkrug,S.C., Frickling,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keeler,J.W.
TITLE	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL	Unpublished (2000)
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel.: 402 762 4366 Fax: 402 762 4390 Email: smithemal@marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -mismatch 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCCTCCAGTCACGAG Plate: 90 row: J column: 21 Seq primer: ATTGTGGTGACATTTAG. Location/Qualifiers 1..506 /organism="Sus scrofa" /db_xref="taxon:9823" /clone_lib="MARC 1PIG" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."
BASE COUNT	91 a 180 c 158 g 77 t
ORIGIN	
Query Match	0.5%; Score 46.4; DB 10; Length 506;
Best Local Similarity	45.7%; Pred. No. 0.65;
Matches 161:	Conservative 0; Mismatches 191; Indels 0; Gaps 0;
QY 1458	ccgcgctcgagaggagacagcgagagcgtcggtctatctcgcgcccttgacaagc 1517
DB 144	CCTAAGAATTGTGAACAACGCGTAGGCCCTCATGACACTGTGAAGTCGAGCTGACCGGG 203
QY 1518	cgcctcacagcgtcgctgttcaccggacagagcgatccgcgtgctcccaaccacccttagcg 1577
DB 204	CAGCCCAAGCCCTGCTGTGTGACCGAGACTTGTCCGAGGCTTGGCATGACCTCGAGCTACT 263
QY 1578	agaagcagcgtgtgctgcgtctgtttcaagtctcgtgctcggagaagcgcggtctgatccagcc 1637
DB 264	TGTTCCGGAGCGCGGCACATCACTAACCTACCCATTGTGAGAAAGGCGCCGTGAGCCGCGCT 323
QY 1638	ttaagagctgtaactcgctcgcgcggaaccttcccagccaactctaacgcgagagacgga 1697
DB 324	TCCGGGGGAGAGACGCGCTGCGCCGGTACCATCCGGGGAGAGACGCTTGCATCGCTTGA 383
QY 1698	atctcgcgcgccagacagcgtcgctcggtgycalcgtccagcatctgagagcaccaaggagtaca 1757
DB 384	AGCTTTGCGAGGCGCTGTCCCTGTGCCACAGCCATCACATCGAGGCGGAGCGGAGGCGCG 443
QY 1758	tcggcaccagagcaagcgccctcgatatctgcacaggaagtatacctcta 1809







Db 375 AGCTTTGGAGGCGCTGTGCTTCCTCCAGGCGCATCACCATCTAGAGCGCCAGCAGGCGCG 434

Qy 1758 tcggacccgagcaagccgagagccctcgatatacgacggaggtatcacctcta 1809  
||||| - - - - - ||||| - - - - - |||||  
Db 435 ACGGAGCGCGCGGACCGCGCTATGACATCGACATGACCAAGTCATCTTA 486

RESULT	15
LOCUS	AM784766
DEFINITION	558 bp mRNA linear EST 09-JUL-2000
ACCESSION	U15012 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
VERSION	AM784766
KEYWORDS	AM784766.1 GI:7841542
SOURCE	EST.
	pig.

ORGANISM	REFERENCE
<i>Sus scrofa</i>	Fahnenkrug, S.C., Fickling, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.
Euxariptota, Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;	
Mammalia, Eutheria; Cetartiodactyla: Suina; Suidae: Sus.	
1 (basses 1 to 558)	

TITLE	Design and use of two pooled tissue normalized cDNA libraries for
JOURNAL	EST discovery in swine
COMMENT	Unpublished (2000)
	Contact: Smith TPL

Single pass sequencing. Bases called and alt-trimmed with phred v0.980904-e. Vector identified by cross\_match with the -minscore 18 and -mismatch 12 options.

FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCCAGTCACGACG  
 Plate: 43 row: M column: 13  
 Seq primer: ATTTAGGTGACACTATAG.

FEATURES	Location/Qualifiers
source	1. .558

```

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT
ORIGIN
99 a      200 c      175 g      84 t

```

Query Match	0.58	Score	46.4	DB	9	Length	558
Best Local Similarity	45.78	Pred. No.	0.63				
Matches	161	Conservative	0	Mismatches	191	Indels	0
						Gaps	0

Accession	Sequence	Position
Oy 1458	ccgcgcgttcggaaggagacaggaagagctcgcggctctatctcgcggccctctaaacg	1517
Oy 127	CCTCAAGATTGTTGTAACATCCTGTGACCCCTCGATGAGACTGAAGTCGGTGACTGTACCGGG	186
Oy 1518	cgctgcacagctgctgctgtttccaccggaacaggaatccgcgtcgtgtccacacccaactttagc	1577
Db 187	CAGCCAGACCCCTGTCGTGAGACGACTGTTGCCAGAGCCTGGGCAATGACCCGAGACTACT	246
Oy 1578	agagcagcgtcgttcgacgttctaaagtctctgctgcgagaagcgcgcttcgatcccaagc	1637
Db 247	TGTTCCGGGAGCCGGCCACCATCATATACCATTTGAGAAAGGGCCGCTGTAGCCCGCGCT	306
Oy 1638	tctacggagctgcgaactgcgtcgcgcgaacacttcccacagacacatctacacgcgagagaccgga	1697
Db 307	TCCCGGGGAGACAGCCCTCGCCCGGATACCATCCGGGAGGAGCGCTTGATCGGCTTCA	366

[illegible]

```
Search completed: August 20, 2002, 09:12:11
Job time: 10889 sec
```



---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 19:01:33 ; Search time 3099.9 seconds  
(without alignments)  
16991.546 Million cell updates/sec

Title: US-09-826-206-4  
Perfect score: 2517  
Sequence: 1 ggcaatggctgcgaattcat.....tgaggaataatagggaacgtg 2517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBml:\*

1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_com:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlgc\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query  
No. Score Match Length DB ID Description

1	2517	100.0	2517	6	AX281576	Sequence
2	2517	100.0	8509	6	AX281574	Sequence
3	2517	100.0	8509	6	AX281868	Sequence
4	2007	79.7	2112	6	AX281573	Sequence
5	2007	79.7	5859	6	AX281575	Sequence
6	49.2	2.0	81	6	AX019311	Sequence
7	47.4	1.9	125020	9	AF429315	Sequence
8	46.2	1.8	1969	9	AF429315	Sequence
9	46.2	1.8	1969	6	AX024376	Sequence
10	45.8	1.8	7218	6	AX024269	Sequence
11	45.6	1.8	32784	6	CE00865	Sequence
12	44.2	1.8	125020	9	AF429315	Sequence
13	43.2	1.7	3490	4	BF414556	Sequence
14	43	1.7	204917	9	AC008040	Sequence
15	42.8	1.7	13291	9	AE005932	Sequence
16	42.4	1.7	1141	6	AX083744	Sequence
17	42	1.7	2325	3	HER278115	Sequence
18	41.6	1.7	109489	2	AL669833	Sequence
19	41.6	1.7	112944	9	AC025761	Sequence
20	41.6	1.7	191035	9	AC093527	Sequence
21	41	1.6	16488	1	AE004992	Sequence
22	41	1.6	44364	3	CE0167D2	Sequence
23	40.8	1.6	10223	1	AE005476	Sequence
24	40.8	1.6	144079	2	AC024342	Sequence
25	40.8	1.6	170216	9	AC008432	Sequence
26	40.8	1.6	296827	1	AP002561	Sequence
27	40.4	1.6	731	6	E03296	Sequence
28	40.4	1.6	11584	1	AE000333	Sequence
29	40.4	1.6	16969	1	D90875	Sequence
30	40.4	1.6	74183	2	AC105624	Sequence
31	40.2	1.6	81	6	AX019312	Sequence
32	40.2	1.6	62192	9	AP003033	Sequence
33	40.2	1.6	174155	2	AC090834	Sequence
34	40	1.6	180379	2	AC106455	Sequence
35	39.8	1.6	10886	6	AX347036	Sequence
36	39.8	1.6	138020	9	AC024776	Sequence
37	39.8	1.6	152861	9	AC104779	Sequence
38	39.8	1.6	165147	2	AC021342	Sequence
39	39.8	1.6	309026	2	AC006760	Sequence
40	39.6	1.6	41322	3	CE050D6	Sequence
41	39.4	1.6	1141	6	AX083744	Sequence
42	39.4	1.6	184410	2	AC106470	Sequence
43	39.2	1.6	10940	8	SLY224356	Sequence
44	39.2	1.6	58067	2	AC102933	Sequence
45	39.2	1.6	67870	2	AC083787	Sequence

ALIGNMENTS

RESULT 1

AX281576 LOCUS AX281576 2517 bp DNA linear PAT 02-NOV-2001

DEFINITION Sequence 4 from Patent WO0177347.

ACCESSION AX281576

VERSION AX281576.1 GI:16608828

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

synthetic construct

artificial sequence.

REFERENCE

1 (sites)

D'Elia, J.

Patent: WO 0177347-A 4 18-Oct-2001;

Archer-Daniels-Midland Company (US) ; D'Elia, John (US)

TITLE

Journal

Location/Qualifiers

1. .2517

location="synthetic construct"

/db\_xref="taxon:32630"

/note="ketogulonigenium part of pADM291-4"

BASE COUNT 657 a 589 c 683 g 588 t

ORIGIN



Query Match 100.0%; Score 2517; DB 6; Length 2517;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggcatagtgctgaatacctatagaatttggtagagtgctgagcgctctgacagaggtg 60  
Db 1 GGCATAAGGTCGAATAATCATAGAAATTTGTGAGGTGGTAGCGGCTGTGACAGGGGTG 60  
Qy 61 ctgcgcggaatctcttggtctcaggtagagggcgaacatggagagtgtagtgcctctg 120  
Db 61 CTGCGCGGAATCTCTGGTCTCAGTGAAGGCGACAAATGGAGAGGTGTAGTGGCCCTG 120  
Qy 121 tatcgctctcgtgagcgatggtagtgcacgtccgcgacataatgatcttcgcctagag 180  
Db 121 TATCGCTCTCGCTGGCGCATTTGGTCATTCGCCGCGCATATATATTCGCTAGAG 180  
Qy 181 gattaactgatatgttctgcgcgtgctggagcttgcgagcttgcgagcttgcgagcttgc 240  
Db 181 GATTACTGATAGTTTCTGCTGCTGCGGCTGTGCGGCTTGTGCGGCTTGTGCGGCTTGTG 240  
Qy 241 gggcctgctccctctgtgccgcgtgctccacatttccacaatcaaaaatggcggaagc 300  
Db 241 GGGCCTGCTCCCTCTGTCCCGCTGCTCCTCACTTTTTCACAAATCAAAAATGGCGGAGC 300  
Qy 301 cctctctgttctatagttctctatagttcaacgaataacacataaattacacatgct 360  
Db 301 CCTCTGTGTTCTATAGTTCTTATAGTTTATGATTCAGAAATTTACACATATATTCATATCCT 360  
Qy 361 attcgcttaaaaggagtaattggcgcaaaaggagtaattggcgcaaaaggagagag 420  
Db 361 ATTTCGCTTAAAGGAGATTAATGGCGCCGCAAAAGGAGTAATTTGGCCGCGCAAAAGGAGT 420  
Qy 421 aattggcgcaaaaggagtaattggcgcaaaaggagtaattggcgcaaaaggagag 480  
Db 421 AATTGGCGCGCAAAAGGAGTAATTTGGCGCCGCAAAAGGAGTAATTTGGCCGCGCAAAAGGAGT 480  
Qy 481 ccttaactatcttccccaatgggaagacacacaaagtgccgagacccgagccttgcag 540  
Db 481 CTTTAAATCATTTCTCCCATGGGAAGACAAACAAAGTGGCGGAGCCGCGCTTCCAGC 540  
Qy 541 caagcaaaaactgtgtccctgcgcgaggtgcaagaggggtctatatagtcgaatccgccc 600  
Db 541 CAGCAAAAACCTGTGCTCCTGCGAGGTGCGCAGAGGGGTCTATATGCGCAATCCGCC 600  
Qy 601 cgctgcagcgctcaagctcaltgcatltaatgataagcaactgcgagcgcgagcgttggt 660  
Db 601 CGCCTGCAGGCGCTCAAGCTCATGCAATTTATGATAGCCACTGCGGCGCGCGCATGCGCT 660  
Qy 661 gatgatgtgcacatgaagaatggggtctggcgacatcgcgcaatcgacgcatganaaac 720  
Db 661 GATGATGTGCGCCATGAAATGGGCTGGCCGACATTTGCGCAMAATGCGGATGANAAC 720  
Qy 721 catgacgtgagagcctgagcccgctgttcgagagagctagccgctgaggtgttgagccat 780  
Db 721 CATGACGTGAGAGCCTGACCCCGCTGTGCGAGAGCTTACCCGCTGGGTGTGACCCAT 780  
Qy 781 gatgacctgcaagaatgatctgtagacgtcgcgcttggtgcgaaggagagagagag 840  
Db 781 GATGACCTTCGCAAGATATGATGACAGTTCGGCGCTTGGTGCATGAGGCGCGCAATGAGAC 840  
Qy 841 taacgcgaagagcaagcgagcaactcctagtgcgtgagacgttcgcgagagagagag 900  
Db 841 TACCGCGAGAGCAAGCGGCAACTCTAGTGAAGTGAAGCTTCCGAGTACATTCGCT 900  
Qy 901 cgtatggcgaggtgcgaacacactgggagcattcgcgacgctcaaacgtaattcacatctc 960  
Db 901 CGTATGGCGAGGTGCGAACACTGGGAGCATTTCGACACGTCANAAAGTATTCATCTC 960  
Qy 961 ggtgaagagatcctgctgtctgtctcagcagcgtctctagtctcgcgaactctgagtg 1020  
Db 961 GGTGAAGATCTCGTGTCTGTCTCCAGCAGCTCTAGTCTGCCAATCTTATGTCGG 1020

Qy 1021 atgagcgcaaaaaccccttaacgttcccgaggttgcgcgctccctcttgagatgccagagga 1080  
Db 1021 ATGAGCGCAAAAACCTTTCAGGATCCCGAGCTGGCGGCTCTTGTGAGTCCCGAGGGA 1080  
Qy 1081 aagatgttcgttggaacagcttacaagatttgtctcaaacctgcacgtgagatgc 1140  
Db 1081 AAGATGTTCTGTTGGAACAGCAGTATACAGATTGTGCTCAAACTGCACTGATGATGATC 1140  
Qy 1141 aacattatcgcgtctgacatctgacgcgaagacggaacaaattggccgtgagctgagca 1200  
Db 1141 AACATTATTCGGGTCTGACATTTGACGGCAAAAGCCGACAAAGTTGGCCGTAGCGTGGCA 1200  
Qy 1201 agtgtgctataagctctgagagagtgaaagacgacaccgctgcagcgagcgagctgagc 1260  
Db 1201 AGTGTGCTATAGCTTGGAGTGAAGTGAAGACGCCAACGCTGCGCAGCGCATGGGCG 1260  
Qy 1261 ggtcccaaggtcgtgtagatgtcgttcgagaggggagcgagcaaacgataagccccctcc 1320  
Db 1261 GGTTCCAAGGTGCGGTCTCAGATGCTGTCGCAAGGGGCGACCGGAAACGATAGCCCCCTCC 1320  
Qy 1321 ttcccaagagcgggcgagatcacctacagtcacagctgtgctgtagagctgaaagcgtcgt 1380  
Db 1321 TTCCCAAGAGCGGGCGGATCACCTACAGTCCAGCTTGCGTGGAGCTGAACGCTGTGCT 1380  
Qy 1381 ggcagcaacaaaggacacagatcgtatgcctcagacttcccgagcttctgctcgagagga 1440  
Db 1381 GGCAGCAACAAAGGACACAGATCTGATTCGCTCAACCTTCGCGCTTCTGTGCGGAGAGA 1440  
Qy 1441 ggcgtgagctgtgagcgtcgaacacacacgaacaaactgttttagatctctgcgcaagta 1500  
Db 1441 GCGTGAGCTGTGAGCCTGACCGCTCCAAACATTCGAAACATGTTTATGATTTCTGCGCAAGGTA 1500  
Qy 1501 gggagagcttgagtttgaggtatctcaacggaatagtggttaattgactcttcgtagaacg 1560  
Db 1501 GGGAGAGTTTGAATTTGAGTATTTCAACCGCAATAGTGTTAATGACTTTTCGTGAAACG 1560  
Qy 1561 atgtgcaatataagcggtaagactatgaatatcacagcgctgagacagctgcaaaagcaagc 1620  
Db 1561 ATGTGCAATATAGCGGTAAAGACTATGAATACACGCGCTGAGACGCTGCAAAAGCAAGCG 1620  
Qy 1621 gtttgagcagcgcaacatcaactcgggctcaaaaagcggtaaatttcgcgttaaaaag 1680  
Db 1621 GTTGTGCGAGCGCAACCATCACTCGGGCGCTTAAAGCGGTAAATTTCCGCTTAAAAAAG 1680  
Qy 1681 atgatatgggagatgggtatagatccctgcagaaatttgacagagtggttctcccatctt 1740  
Db 1681 ATGATCTGGGGCTGGGTTATGATCTGCAAAATTTGCAACAGAGTGTTCCTCCCATTT 1740  
Qy 1741 caagaaatacacccgaaacacactaacagcgaagtatatggtgaaggtgtagaacaacatg 1800  
Db 1741 CAAGAAATACACCGAAACACCTTAAACGCAAGCATATATGCTGATGAACACCATG 1800  
Qy 1801 aaatgacctgaagaatcagcgcatctagagctggaagttcgagacttcaagatgtcttat 1860  
Db 1801 AAATGACCTCGAANAATCGAGCATTTAGAGGCTGAAGTTGCGACTTTACGGAGTCTTAT 1860  
Qy 1861 ctgattgcagagagagatcgcgacaatggcgagcatagcgcgagcgctcttcaattcat 1920  
Db 1861 CTGATGCCAGGAGGATGCGCAAAATGGCGCATGAGCGGACGCTTTCAATTTTCAT 1920  
Qy 1921 caacgatagagagagagacgccccctcaaaaacaaagatggtgtagaataattctgat 1980  
Db 1921 CACCGATGAGAGAGAGACCGCCCTCAAAAACAAAGATGCTGAGATATTTGAT 1980  
Qy 1981 cctggagctcaagagccttgcttaaaactggaatcaagcattctagcgtatgctgataag 2040  
Db 1981 CTTGGGCTTCAGAGGCTTGTCTTAAACCTGAATTCAGCATTTCTAGCGATGCTGATAG 2040  
Qy 2041 aagtaatatagcgaacaaatagagcgccatttccattcaatacagctcaatctatgagat 2100  
Db 2041 AAGTAAATATAGCGCAATATAGAGCGGCGCTTTTCCATTACATACAGCTCATATGATGAT 2100  
Qy 2101 caatatcaagtatgtatattcatlcaatggagaagaattacatgatacaagagatcatca 2160



|||||  
Db 2101 CAATATCAAGTATGATATCATCATGAGAGAAATTACATGATATCACAGATCATCA 2160  
Qy 2161 cagatattgttttgatttctaagtgtaataactatcatgcctgagccctaagaagttc 2220  
Db 2161 CAGCATTTGTTTTGTATTTCTTAAGTGTACATATACATATGCTGTGCCCCATAAGAAATY 2220  
Qy 2221 gtactatgcagatcacccttgggttcagaccgaataatcagcttctcgaagccg 2280  
Db 2221 GTACTATTGCAGTATCTACCTTGGGTTTCAGACCGATATTTACAGCTTTGTGCGAAGCCG 2280  
Qy 2281 gttttttgcagagagagacagcttcttgatgltglaataaactgctacgtatcctc 2340  
Db 2281 GTTTTTTGGCAGAGAGAGACACGTTTTTGTATGTGTATTAACATGCTACGATCTCATAG 2340  
Qy 2341 atggtataataacacagatcatccggggcaacacacacctcttatagtgaagtatctg 2400  
Db 2341 ATGTATACATACACAGCATCTCCGGGGCAACACACTCTTATGTGAAGATGATATTTATG 2400  
Qy 2401 gcccaagaactcgcgcgcaaaaacgcgatatgtgaagcagagcccgacactttaagagcg 2460  
Db 2401 GCCCAGAGTACTGGCGGAAAAACCGCATATTGAGGCAACAGCCCGCCTTTAGAGCGCG 2460  
Qy 2461 aagcctataacgagtlaccaaaacactagaagccagattcgaagaaatagggaaacgtg 2517  
Db 2461 AAGCCTATAACGAGTACCAAAAACACTAGAACCCAGATTGAGAAATATGAGAAACGTG 2517  
RESULT 2  
AX281574 8509 bp DNA linear PAT 02-NOV-2001  
LOCUS AX281574  
DEFINITION Sequence 2 from Patent WO0177347.  
ACCESSION AX281574  
VERSION AX281574.1 GI:16608826  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequence.  
REFERENCE  
1 (sites)  
AUTHORS D'Elia, J.  
TITLE Ketogulonigenium shuttle vectors  
JOURNLS Patent: WO 0177347-A 2 18-Oct-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
source  
1. 8509  
/organism="synthetic Construct"  
/db\_xref="taxon:32630"  
/note="PADM291"  
BASE COUNT 2065 a 2239 c 2345 g 1860 t  
ORIGIN

|||||  
Db 3195 GGGCGTTCCTCTTGTCGCCGCTCTCCATCTTTTTCACATCAAAAATATGGCGGAAGC 3254  
Qy 301 cctcttgctctatagtttcttaagttcataagaaaatlaacataatattcaatagctt 360  
Db 3255 CCTTTCTTGTCTATATAGTTCTTATATGTTATACGAAATATATACATATATATATATAGCTT 3314  
Qy 361 attcgcttaaaaggaggaatattggccgcaaaaggaggaatattggccgcaaaaggagct 420  
Db 3315 ATTCCCTTAAAGGAGTATTTGGCCCGCAAAAGGAGTATTTGGCCCGCAAAAGGAGCT 3374  
Qy 421 aattggccgcaaaaggaggaatattggccgcaaaaggaggaatattggccgcaaaaggagct 480  
Db 3375 AATTGGCCCGCAAAAGGAGTATTTGGCCCGCATATTCGTTGTTACATGSGGAGGATTC 3434  
Qy 481 ccttaataattctcccataagggaagacacacaagtgccgcgaagccggccttcgac 540  
Db 3435 CCTTATATATTTCTCCCTCCATGCGAAAGACMACAAAGTGGCCGACAGACCGGACCTTCGAC 3494  
Qy 541 cagacaaaactgtgctccctgcgaggttggcgagaggggtctatatgccaatccgccc 600  
Db 3495 CAGACAAAACCTGTCTCCCTCCGAGGTGGCGAGAGGGGTCTATATGCGCAATCCGCC 3554  
Qy 601 cgcctgcagcgctcaagctcatalgcatltaatagccaactgcgagcgagccgcaatgct 660  
Db 3555 CGCCTGCAGGCGCTCAAGCTCATGCAATTATATAGTACGACATGCGGGCGCCGATGAGCT 3614  
Qy 661 gatgatgtgcgcataagaaatgctgctgcgcgacatctgcgcaatcgaagcatgaaaac 720  
Db 3615 GATGATGTGCGGCATGAATATGGCGTGGCGCATTTCCGCAATGACGCGATGAAAAAC 3674  
Qy 721 catgaacgtaagagccctgcagcccgctgttcgagggagctgcgcgtgcgtgttgcagccat 780  
Db 3675 CATGACCGGTGAGAGCTGACCCCGCTGTTCAGAGACCTAGCCGCTGCGGTGTTGACCCAT 3734  
Qy 781 gatgacctgcacaagaatgacgtgacagtcgagcgagcttgatcgaatgagcgcgaaatagac 840  
Db 3735 GATGACCCCTGCAAAAGATGATGCTGACAGTGGCGGCTTGTGATGAGCGCGAATATGAC 3794  
Qy 841 tacgcgcagagaggaagcgagcggaactcctagtgaagtggaacttcggaagtaactcgt 900  
Db 3795 TACCCGCCAGAGGACGAGCGCGCAACCTCTATGAGCTGAGACCTTCCGGATCATTCCT 3854  
Qy 901 cgtaaggcgagagtgagacacacttgggacattctcgagccgtcaaacggtattcacttc 960  
Db 3855 CGTATGGCGCGGAGTCGAACCACTGGCCATTCTCGACCGTCAAAAGGTATTCATCTC 3914  
Qy 961 ggtagtaagtaattcgtgctgctgtgttcacagacgctcactagctcgcgaactctgacg 1020  
Db 3915 GGTAGTAAATATTCGCTGCTGCTGCTTCAGCAAGCTCTAGTCTGCGCAATCTTGATTCGG 3974  
Qy 1021 atgagcgcaaaaacctttaggttcccgagattggcgagctcccttgaggtgcccgagga 1080  
Db 3975 ATGACGCGCAAAACCTTTACGTGCTCCCGAGTGGCGGCGCTCTTGAGTGGCCGAGGCA 4034  
Qy 1081 aagatggttcgtttggaacgagcttaacagattgctctcaaacctgcagcttgatgagtc 1140  
Db 4035 AAGATGTGTTCTTTGGAACGACGTTAACAGATTTGCTCTAAACCTGCACTGATGATGATC 4094  
Qy 1141 aaccaattacgctgctgacatctgacgcaaaagcgcaaacgaattgcccgtatgtagca 1200  
Db 4095 AACCAATTTATCCGCTGTGACATTGACGGCAAAAGCCGACCAAAATTTGGCGGTAGCGGCA 4154  
Qy 1201 agtggactatagctcgtggaagaattgaaagcgacccaacgcttgcgaagcgagagctgg 1260  
Db 4155 AGTGTGACTATATAGCGCTGGGAAATGAAAGACCCAAACCTGTGCGGCGGCAAGCTGGG 4214  
Qy 1261 ggtccaaagtcgctcgaatgctcgtcgcagaagggcgagaaacgataagccctcc 1320  
Db 4215 GGTTCCAAGTGTGCTGAGATGCTCGTCCAGAGGGGCAAGCGGAAACGATACCCCTCTCC 4274  
Qy 1321 ttcccagaagcgagcgagatcaactacatcagttcagcttgagagctgaaacgctctgct 1380  
|||||



Db 4275 TTCCCAGAAAGCGGGCGGATCACCTTACAGTCCACGTTGGCTGGAGCTGMAAGCTCTGCT 4334  
QY 1381 ggcagacaagaagacaagcatcgtatcgccctagactccgcgcttctctgtcgtgaagaga 1440  
Db 4335 GCGAGCAACAAGACAACGATCGATCGCCTCAGACTTCGGCGTTTCTGTCGGAGAGA 4394  
QY 1441 ggcgtgcgtctgtagacgtcacaacatcgaaaaacgtcttctagattctctgcgaagta 1500  
Db 4395 GGGGTGCGTGTGAGCGCTGCAACATCGAAAACTGTTTTTATGATTCTCGCCAAAGTA 4454  
QY 1501 gggaaagtttgaagttttgaagtaattcaaccgaaatagttttaagacttctgtgaagc 1560  
Db 4455 GCGAAGGTTTGAGTTTGAGAGTATTCACCGCAATAGTGTAAATGACTTTCGTGAAGC 4514  
QY 1561 atgtgcaataatagcgttaaaactatgaataacacgtctgacagcgtcgaagaacagc 1620  
Db 4515 ATGTGCAATATAGCGGTAAGACTATGAAATACAGGCTGACAGCGCTGCAAAAGCAACG 4574  
QY 1621 gttgtggagaccgcaacatcactcgtgcgttaaaaaagcgttaaaattccggtlaaaaag 1680  
Db 4575 GTGTGGCCACGCAACCATCTCGGGCGCTAAAAAGCGGTAAATTTCCGTTAAAAAAG 4634  
QY 1681 atgaacttggggcatgggttatatagatccctgcaaatctgcacagatggttccctcccaatt 1740  
Db 4635 ATGAATCTGGGGCATGGCTTATAGATCCTGCAGAAATTCACAGAGTGTTCCTCCCATTT 4694  
QY 1741 caaagaaatacagcaaaacaccttaacacgcacgcatatgtaagcgtatgaacacatg 1800  
Db 4695 CAAAGAAATACCCCAAAACACTTAACAGCAAGTATATGTAAGCGTATGAACACATG 4754  
QY 1801 aatgacctcagaatacagcgcatlagagcgttgaagttcgacttctgcagatgcttat 1860  
Db 4755 AATGACCTCAGAAATTCAGCGCATTCAGCGTGAAGTTCGGACTTTACGCATCTTTAT 4814  
QY 1861 ctgtagcagagagagatcgcagcaaatgacgcagatgacgcgacttctcaattcaat 1920  
Db 4815 CTGATGCCAGGAGAGATCCCGCAAAATGGCGAGATGGCGAGCGTCTTCAATTCAT 4874  
QY 1921 caacgtagaagagaagaacgcgcgcctcaaaaaaagaatggtggaagaattctgcat 1980  
Db 4875 CACCGATGAGAGAGAGAACCCGCCCTCAAAAACAAAGATGGAGAGATTTCTAT 4934  
QY 1981 cctggcctcagaagccttgcctttaaaacctgaatcagcatctcagatgctgataag 2040  
Db 4935 CCGGGCTTCAGAGCCTTGCTTTAAACCTGAATCAGCATTTAGCGATCTGATAG 4994  
QY 2041 aagtaataatagcacaatagagcgcaatttccatcaatacagctcatcatgcat 2100  
Db 4995 AAGTAATATAGCCCAATAGAGCGGCCATTTTCCATTACATACGCTCATCATGTGAT 5054  
QY 2101 caataatcagatgatattcatcaatgtgagaagaattacatgtaatacagaatcatca 2160  
Db 5055 CAATATCAATATGATATTCATCAATGAGAGAAATTTACATGATTCACAGAGATATCA 5114  
QY 2161 cagcatgttttctgtaattcttaagtctaaactaaactcgtgcgccttaagaagatt 2220  
Db 5115 CAGCATTTGTTTGTGTTCTTAAGTCTAACATTAATGCTGCGCCCTAAAGAGAT 5174  
QY 2221 gtaactatcagtaatacactcgttggtttcagaacgataatatacacttctgtcgaagcg 2280  
Db 5175 GTACTATTGCAAGTATCTACCTTGGTTTACAGCCGATTAATTAAGTTTGTCGAAGCCG 5234  
QY 2281 gttttttgcaagagagacagtttttgatggtgtaataaactgtaactgtaactcga 2340  
Db 5235 GTTTTGTGCGAGAGAGACACTTTTGTGATGTTGTAATAAATGCTAGTATCTCATG 5294  
QY 2341 atgtgaataatacagcatcatccggggcaacacactctatggaagatggatattatg 2400  
Db 5295 ATGTAAATACACAGCATCATCCGGGGCAACACACCTTTATGGAAGATGATATATG 5354  
QY 2401 gcccaagaagctcgcgaaaaaagcgaatctgagcgaagcccgactttaagcgcg 2460  
Db 5355 GCCCAGAACTACTGGCGGAAAGACCGCATATTTAGAGGACAGGCCCGCATTTAGAGCGG 5414

QY 2461 aagcctataacagtgatccaaaacactagaagccaatttggagaaataaggaactg 2517  
Db 5415 AAGCCTATTAACGAGTACCAAAACACTAGAACCCAGATTGAGGAAATATAGCAACGTG 5471

RESULT 3  
AX281868 8509 bp DNA Linear PAT 02-NOV-2001  
LOCUS  
DEFINITION  
Sequence 1 from Patent WO0177159.  
ACCESSION  
AX281868  
VERSION  
AX281868.1 GI:16609119  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified.  
unclassified.  
unclassified.  
REFERENCE  
1 (sites)  
AUTHORS  
Schmidt,T.M. and Stoddard,S.F.  
TITLE  
An endogenous ketogulonigenium plasmid  
JOURNAL  
Patent: WO 0177159-A 1 18-OCT-2001;  
MICHIGAN STATE UNIVERSITY (US) ; Archer-Daniels-Midland Company (US)

FEATURES  
source  
Location/Qualifiers  
1..8509  
/organism="unidentified"  
/db\_xref="taxon:3264"  
/note="ketogulonigenium"

BASE COUNT 2065 a 2239 c 2345 g 1860 t  
ORIGIN

Query Match 100.0%; Score 2517; DB 6; Length 8509;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggcagaggtcgaatctcaatagaatttctgtgagtgctagcgcgtctgcagaagggtg 60  
Db 2955 GGCATAGGGTTCGAAATTCATAGATTTTGTGAGGTGCGTAGCGGCTTCACAGGGGTG 3014  
QY 61 ctgcgcggaagatctctgtctcagtagagcgacaattggaagagtgtagtgccctg 120  
Db 3015 CTGCGCGGAGATCTCGTCTCAGTAGGGCCGCAATGAGAGGTGATGTTGCTCCCTG 3074  
QY 121 tatgcctctgtgtgagcagatgggtcaatcctgcgcggaacatagatattccgtagag 180  
Db 3075 TATGCTCTCTGCGTGGCGATTTGGGATCCGCGGACATATGATTCGCTAGAG 3134  
QY 181 gattactagatgttctgcgtctgcggtctgtcgcgctgtcgtggtgctgtggtgctgtgc 240  
Db 3135 GATTACGATAGTTTCTGCTGTCGGCTTGTGGGCTTGTGGGCTTGTGGGCTTGTG 3194  
QY 241 gggcctgtccctcttgcgcgcgctgtcactcttccacaatcaaaaaatggcggaagc 300  
Db 3195 GGGCTGTCCCTCTGTCGCCCTGCTGCTCACTTTTCAAAATCAAAAAATGGCGCAAG 3254  
QY 301 ccttctgtctatagttctatagttcattacatcgaanaatacaataatcaatagctt 360  
Db 3255 CCTTCTGTTCTATAGTTCTTATAGTTCATACGAAATTCACATTAATTCATAGCTT 3314  
QY 361 atcgccttaaaagggaatattggcgcaaaaaggagtaatttggcgcgaagaaggagt 420  
Db 3315 ATTGCTTTAAAGGAGATTAATTTGGCCGCAAAAGGAGTAATTTGGCCGCAAAAGGAGT 3374  
QY 421 aattggcgcaaaaagggaatattggcgcatatcggttcttatacattgggaagatcc 480  
Db 3375 AATTGGCCGCAAAAGGAGTATTTGGCCGATATCGGTTGTTTACATGGGAGGAATTC 3434  
QY 481 ccttaatcatcttccccaatgggaagaacaacaagtggcggaagccggccttcgac 540  
Db 3435 CCTTATCATTTCTCCCATGGAAGAACACACAAAGTGGCGGACCGGCTTGCAC 3494  
QY 541 cagacaaaacgtgtccctcctgcagagtggtggaagggtctatatatgcaatccgcc 600



D	b	3495	CAGACAAAACACTGTCCTCCCTGCCGAGGTGGCAGAGGSGGTCTATATCCGAATCCGCC	3554
Q	y	601	cgccctcgacaggcgtctcaagctcatalgtcaattatagatagcaactctgcgcgcgcgaatgct	660
D	b	3555	CGCCTGACAGGCGGTCAAGGTCATGACATTTAAATGATATGCACATCGGGGGCGCATGGCT	3614
Q	y	661	gataatgtctgcgcacatgaaatctgcgtctgcgcgaatctgcgcaatctgcacgcaatgaanaac	720
D	b	3615	GATGATGTGGCGCCATGAAAATGCGGGCTGGCCGACATTCGCGCAATGCAGCGCATGAAAAAC	3674
Q	y	721	catbaacgttaagagccgtgaccgcccgctgtctcgaggagctagccgcgtgcgtgttgaccat	780
D	b	3675	CATTAACCGTGAAGCCCTGACCCCGGCTTCGAGAGACATGACCGCGTGGCTTGAACCAT	3734
Q	y	781	gatacccttcaaaagatgatactgtaacagtcgcgcgcgtctgtctgaatgaagcgcgaatagac	840
D	b	3735	GATACCCCTGCAAAAGATGATCGTGAACATGTCGGCGGCTTGATGATGAGCCGGAATAGAC	3794
Q	y	841	taaccgcagaagagcaagcgcgcgaactcctcagtagtcgtgacctctccgaatatactcgt	900
D	b	3795	TACCGCAGAGGCGCAACGCGCGCAACTCTTAATGATAGCTGAGCCTTCGGAGATCAATCCGT	3854
Q	y	901	cgatctgacgcgcgagtcgtcaaacccactctggccattcttcgaacgtcaaaagtagttccatctc	960
D	b	3855	CGATATGCGCGCGGAGTGTGAACCACTGGCGCATTTCTGACCGTCAACCGGTATTCATCTC	3914
Q	y	961	ggtatgaatgattcttcgcgtctgccttccaaagctctcagctctgcgaacttgatctcg	1020
D	b	3915	GGTATGTAAGTATTCGTCGTCGCTGCTTCAGACAGTCCTTAAGTCTGCGCAATCTTGATCCG	3974
Q	y	1021	atlgacgcgaaaaaccttaccgctcccccagatctgcgcgcgtctcctctgtagtgatcccgagga	1080
D	b	3975	ATGAGCGCGAAAAACCTTTACGGTCCCCGAGTGGCGGGCGCTCTTGAGTGCCTCCGAGGGA	4034
Q	y	1081	aagatggttcgttggaaacgcgttcaacagattgtctcctcaaacctgcacgtgatgagatc	1140
D	b	4035	AAGATGTTGCTTGGAAACGACGTTTAACAATTTGCTCTCAAACTGCACCTGGATGAGATC	4094
Q	y	1141	aaccattatcgcgtctcgtacatatgaaagcgaagcgaaccaagaattgscgtatagcttgcca	1200
D	b	4095	AACCATTTATTCGCTCTGACATTGACGGCAAAACCCGACCAATATGGCGGTAGCGGTGCA	4154
Q	y	1201	agtgtagactaatagctctggaaagatgaaagacgcacccaaacctctgcacagcgcgcagctg	1260
D	b	4155	AGTGTGACTATAGGCTGGAAGTGAAGAAAGACGACCCAGCCGTCGACGCGGAGCGTGGCG	4214
Q	y	1261	ggtctccaaagtctcgatcgatctcgtctgcgcagaagggcagcggaaacgatagtcccccctc	1320
D	b	4215	GGTTTCAAGGTCGGTCTCAGATGCTCTCTCCACAGGGGCGACGCGAAACGATATGCCCCCTCC	4274
Q	y	1321	ttcccagaagcgcgcgcgcgcacactcaagtcacacgtctgactctggaactgaagacgtctgct	1380
D	b	4275	TTCCCAAGAGCGGGCGGGATCACTCAAGTCCACGTTGGCTTGACCTGAACAGCTCTGCT	4334
Q	y	1381	ggcagcaacaagagcaacgcgtctcgtatcgactcaagatcttcgcgcgtcttctgtcgggaagga	1440
D	b	4335	GGCAGCAACAAGGACCAACGATGTGATTCGCTCAGACTTCGCGCGTTCTGTCGGGAGGA	4394
Q	y	1441	ggcgtctcgtctggaactcgaacaatctgaanaaaactgtttttagattctctgcgcaagta	1500
D	b	4395	GGCGTGGCTGTGGACGCTCGCAAACTCGAAAACTGTTTTGATTTCTGGCGCAAAAGSTA	4454
Q	y	1501	gggaaggtttgaagtttgaagtaattccacgcgaatagtgtaatagattctcgtgaagac	1560
D	b	4455	GGGAAGGTTTGAAGTTTGAAGTATTTTCACCGCAATATGTTTAATGACTTTCGTGTAACG	4514
Q	y	1561	atctgcacaatatagcgttaagaactatgaatatcacgcgtctgcacagcgtcgaagaacgcg	1620
D	b	4515	ATGTGCAATATAGCGGTATAGACATATGAATATACCGGCTGGAACGGCTCGAAAAGCAACGG	4574
Q	y	1621	gtctgacgcaccgcaacatcactctcgcgtctaaagaagcgtgataaaattctcgtgtaanaag	1680
D	b	4575	GTTGGGCGACCGCAACCATCATCTCGGGCGGTAAAAAGCGTAAATTTCCGGTAAAAAG	4634

QY	1681	atgaatctgggggcatggtgattagatctcctgcgaattgcacagagtggttctctccattt	1740
Db	4635	ATGAATCTGGGGCAATGGGTTATATAGATCCCTGCAGAAATTCACAGAGCTGTTCTCCCATTT	4694
QY	1741	caaaagaatatcacccgaaaacacctaaacacgaagtatatgtaagsgtgaatgaacaactg	1800
Db	4695	CAAAGAATATCACCCGAACACCTTAACGCAAGTATATGTGAAGGTGATGAACAACATG	4754
QY	1801	aaatgacctcagaanaatcagcgcatatagagcgtgaagtttcggaacttaacgtagtcttat	1860
Db	4755	AAATGACCTCAGAAATATACAGCGCATTAAGAGCGGTGAAGTTGCGACTTTAAGCGATGCTTTAT	4814
QY	1861	ctgatgcgaaggaggaatcgcgcgaacaaatggcgcgaatggcgcgaagcgtcttcaattcat	1920
Db	4815	CTGATGCGCCGAGGAGGATGCGCAAAATGGCGGACATGGCCGACGCGCTCTTCAATTTGAT	4874
QY	1921	caccgatgagaagaagaagacgcgcgcgcctcaaaaaaacaagaatggtggaagaatattcgat	1980
Db	4875	CACCGATGAGAGAGAGAACGCCGCCCTCTCAAAAAAAGATGTTGGAAGATATTCTGAT	4934
QY	1981	cctgggcttcaggagccttgctcttcaaaacctgaatcagcaltctagcgaatgctgataag	2040
Db	4935	CCTGGGCTTCAGAGACCTTGCTCTTAAAACTGAAATACAGATTTTACGGATGCTGATAG	4994
QY	2041	aagtaataatagccacaataagagcggccatttccatccaataagctcatcatgtgat	2100
Db	4995	AAGTAAATATATACCACCAAAATAGACGGCCATTTTCCATTACATPAACGCTCATCATGTGAT	5054
QY	2101	caatatcaagtatgatattatcaatgaatgaagaatattacatgfatcaagaatcatca	2160
Db	5055	CAATATCAAGTATTTGATATTCATCAATGAGAGAAATTTACATGTAATCACAGATCATCA	5114
QY	2161	cagcaattggttttcttattctaaagtgcataataactacgcttgccctaaagaagatt	2220
Db	5115	CAGCATTTGTTTTGATTTCTTAAGTGTCAATACATATATCGCTGGCCCTAAGAAGATT	5174
QY	2221	gtacctatggcagatctcaaccttgggttctcagaaccgataatctaacagctttgtcgaaagcg	2280
Db	5175	GTACATATTCGCAATATCTCACCTTGGGTTTCAACACCGATTAATTAACGCTTGTGCAAGCG	5234
QY	2281	gtcttttccgagagagacacgcttttgaatggtgtaataaactgcatactatcatg	2340
Db	5235	GTTTTTTTCCGAGAGAGACACGTTTTTGAATGGTGAATTAACGTCTATCTCATG	5294
QY	2341	atggtataacatcacacagcatcatccggyggaacacacctcttatggaagaatgataatg	2400
Db	5295	ATGGTATACATCACACAGCATCATCCGGGCAACACACCTCTTATGGAAGATGATATTATG	5354
QY	2401	gcccggaagtactgycggaanaaocggaatatggaagcccaagcccgccacttaagagcgg	2460
Db	5355	GCCCGAATATCTGGCGGAAAAACCGAATTTGAGGCAACAGCCCGCACTTTAAGGCGG	5414
QY	2461	aagccataacagatgacaaacaactlagaagcagatltgaggaataaataaggaacgtg	2517
Db	5415	AAGCCTATACAGATACCAAAACACTAGAGGCAATTTAGGAAAAATAGGAACGTG	5471
RESULT	4		
LOCUS	AX281573	2112 bp	DNA
DEFINITION	Sequence 1 from Patent WO0177347.		
ACCESSION	AX281573		
VERSION	AX281573.1		
KEYWORDS	GI:16608825		
SOURCE			
ORGANISM	synthetic construct.		
	synthetic construct		
	artificial sequence.		
REFERENCE	1 (sites)		
AUTHORS	D'Elia, J.		
TITLE	Ketogulonigenium shuttle vectors		
JOURNAL	Patent: WO 0177347-A 1 18-OCT-2001;		
	Archier-Daniels-Midland Company (US) ; D'Elia, John (US)		







Db 1981 CCTGGGCTTCAGAGACCTTGCCCTTACTGCGGAA 2015  
|||||  
RESULT 5  
LOCUS AX281575 5859 bp DNA Linear PAT 02-NOV-2001  
DEFINITION Sequence 3 from Patent WO0177347.  
ACCESSION AX281575  
VERSION AX281575.1 GI:16608827  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial construct.  
REFERENCE 1 (sites)  
AUTHORS D'Elia, J.  
TITLE Kelco/Uniongenium shuttle vectors  
JOURNAL Patent: WO 0177347-A 3 18-Oct-2001;  
Archer-Daniels-Midland Company (US) ; D'Elia, John (US)  
FEATURES  
source  
1. 5859  
location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="PADM291-4DS"  
BASE COUNT 1444 a 1487 c 1566 g 1362 t  
ORIGIN  
Query Match 79.7%; Score 2007; DB 6; Length 5859;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggcaatgggtcgaataatcatagaaattgtgtgaggtgctgagggcctgacagggtg 60  
Db 402 GGCATATGGGTGGAATTTATGAATTTGTGTGAGGTGCGTGAAGGCGCTGACAGGGTG 461  
Qy 61 ctgcgcgagatctctgtgtcaggtgagcgacatgagaggtgttagttgcccctg 120  
Db 462 CTGCGCGAGATCCTGTGCTCAGGTAGGGGACATGAGAGGTGTAGTTGCCCTGTG 521  
Qy 121 tatgctctcgtgtgagcatgttgatctatctcgcggaacatatgatatccgctag 180  
Db 522 TATGCGCTCTGCGGGGCAATGGGTCTATCTGCCCGACATATGATATTCGCTAGAG 581  
Qy 181 gattacatagttctcgcgtgaggtgtgtcgggtctgtcgggtctgtcgggtcttc 240  
Db 582 GATTAATATAGTTTCTCTCTGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCTTGTG 641  
Qy 241 gggcctgtccctctgtgccgcgtgtcctacatttccacaatacaaaatgagcgagc 300  
Db 642 GGGCCTGTCCCTCTGTGCCGCTGTCTCACTTTTTCACATCAAAAAATGGCGGAAGC 701  
Qy 301 cctctgttctatagttctctatagttcatagaaatcaacaataatcatagctt 360  
Db 702 CCTCTTGTCTATAGTTCTTATAGTTCTATGCAAAATTACACATATATATCAATAGCTT 761  
Qy 361 attcgcttaaaaggaggaatttggcgcaaaaggagtaatttggcgcaaaaggaggt 420  
Db 762 ATTGCGCTTAAGAGGAGTAATTTGGCCGCCAAGAGGAGTAATTTGGCCGCCAAGAGGAGT 821  
Qy 421 aattggcgcaaaaggaggaatttggcgcaaaaggagtaatttggcgcaaaaggaggt 480  
Db 822 AATTGGCGCGCAAAAGGAGTAATTTGGCCGCCAAGAGGAGTAATTTGGCCGCCAAGAGT 881  
Qy 481 ccttaataatttctcccaatgggaagaacaacaagtggcgcaagaccgggctctgac 540  
Db 882 CCTTAATATTTCTCCATGAGGAGAACACACAAAGTGGCGCAGACCGGGCTTTCGAC 941  
Qy 541 cagacaaanaactgtgtcctcctgaggttggcgagaggtgtatatacgcaatcgccc 600  
Db 942 CAGACAAAACGTGTCTCCTCTCCAGAGGTGGCGAAGGGGTATATATGCGCAATCGCCC 1001  
Qy 601 cgcttcgagcgctcaagctcatgcatltaatgcatagccactgcyggcgccgcatgct 660

Db 1002 CGCCTGACAGGCGCTCAACCTCATGCAATTAATGATAGCCACTGGCGGGCGGCATGTGCT 1061  
Qy 661 gatgatgtgcgcatgaaatgacgtgtgcccgaacatgcgcaatcgaacggaatgaaac 720  
Db 1062 GATGATGTGCGCCCATGAAATGGGTGGCCGATCTCCGCAATGAGCGGCAATGAAAAAC 1121  
Qy 721 catgacgttgagagccctaccccgctgtctgagagagctagccgctgctgttgcacat 780  
Db 1122 CATGACCGTGAGAGCCTTAACCCCGCTGTTCAGAGAGCTAAGCCGCTGCGGTGTGACCAT 1181  
Qy 781 gatgacccctgcaaaagatgcatgtagcaagtcgagcggtctgtgtcgaatgagcgaaatagac 840  
Db 1182 GATGACCCCTGCAAAAGATGATGCTGTGACAGTCGGCGGCTTGTGTGATGAGCGCGGAATAGAC 1241  
Qy 841 taacgcgaagaggaagcgaggaactcctctagtgaagtgagacctcccgagatcatcgt 900  
Db 1242 TACCGCCAGGAGGCAAGGGCGGCAACTCTTAGAGAGTGTGACCTTCGGAATCATTCGCT 1301  
Qy 901 cgtatggcgcgaggtcgaaacactgggccaattctcgaacggtcaaacggtatccatctc 960  
Db 1302 CGTATGGCGGCGAGTCCGAACCACTGGGCCATTCTCGACCGTCAACCGTATTCCATCTC 1361  
Qy 961 ggtatgaaglatccgltgctgtctgtctcagcaagctctctagtctgcgaatcttgatcgy 1020  
Db 1362 GGTATGAGTATTCGCTGCTGCTGTTCACAGACGCTCTAGCTCGCAATCTTGATCGG 1421  
Qy 1021 atgagcgcaaaaccttcaaggtcccgaggtgtggcgctccttgaagtgcggaagga 1080  
Db 1422 ATGACGGGGAAMACCTTTACGGTCCCGAGTTGGCGGCGCTCTTGAGTGGCCAGGGA 1481  
Qy 1081 aagatggtctgttgaagagacttaacagattgtcttcaaacctgcgaactgataagatc 1140  
Db 1482 AAGATGTTCTGTTGGAACGACGCTTAACAGATTGCTCTCAAACTGCACCTGGATAGATC 1541  
Qy 1141 aacacattatcgctgtcgtacattgacgtgcgcaaaagcggaacgaatttgcgttagcgtgga 1200  
Db 1542 AACATTATGCGCGTGTGACATTGACGGCAAGCCGACCAAGATTGGCGGTAGCTGGCA 1601  
Qy 1201 agtgtgactataggtctggaaggtgaagaagaccccaacccgttcgaagcgcgagctgagc 1260  
Db 1602 AGTGTGACTATAGGCTGGGAAGTGAAGACGACCCACCGTGCAGGCGCAGCTGTGCG 1661  
Qy 1261 ggtccaaagctggtcgcgaagatctgctgcgcaagaggcgagcggaagataagccccctcc 1320  
Db 1662 GGTTCGAAGTGTGCTGAGATCTGCTCGCAGAGGGGCAAGCAACATTAACCCCTTC 1721  
Qy 1321 ttcccaagaagcgggcggaatccctacagttccagcttggcttggagctgaaacgctctgct 1380  
Db 1722 TTCCAGAAAGCGGGGATCACCTAAGTCCAGCTTGAGCTGGAAGCTGAACGCTCTGCT 1781  
Qy 1381 ggcagcaacaagaagaagatctgagtcgctcagaagctcccgagcttctgtctggagaaga 1440  
Db 1782 GGCACCAACAAAGGACAGCATGTGATCGCTCAGACTTCGGCGCTTGTGTGGAGAGA 1841  
Qy 1441 ggcgtgctctgagcgctgcgaacaactcgaaaaacggtttttagattcttcgtcgaagta 1500  
Db 1842 GGCGTGCTGTGACGCTGCAACATCGAAAAACGTTTTATGATTTCTGCGCAAAAGTA 1901  
Qy 1501 gggaaagtttgaattcttgaaglatctcaccgaatggtttaaatagacttctgtgaagc 1560  
Db 1902 GGGAAAGTTTGAATTTTGGATTTTACCCGAATAGTGTAAATAGACTTTCGTGAACG 1961  
Qy 1561 atgtgcaataatgcygtaagactatgaataaacaagcttgaagaagcttgaagaagaagc 1620  
Db 1962 ATGTGCATATATGCGGTAAAGCTATGAATTAACAGCGCTGGAAGCTGCAAAAAGCAACGG 2021  
Qy 1621 gttgtgcgacgcgaacacatcactcgagcgctaaagaagtgtaaaatttccggttaaaag 1680  
Db 2022 GTGTGCGACCGCAACATCACTCGGCGCTTAAAAAGCGGTAAATTTCCGTTAAAAAAG 2081  
Qy 1681 atgaatcttggggcatggttataagatcctgcgaagattgcagagtggttctcccatct 1740  
|||||



Accession	Sequence	Position
Dp	2082 ATGAATCTGGGGCATGGGTTATAGATCTCTGCAAAATTGCACAGATGTTTCTCCCAATT	2144
Qy	1741 caagaataatcacccgaataacacttaacacgaaglatatgtaagcgyatataaacacatg	1800
Dp	2142 CAAGAAATACCGCGAAACACCTTAAACCCAGATATATGTAAGGGTATGAACACATG	2201
Qy	1801 aaatgacctaaataatcagcatctatagcgtgtaaglttcgacattaacgagatgcttat	1860
Dp	2202 AAATGACCTCAGAAATCAGGCGATTTAGAGCGTAAATTGGGACTTTACGCGATGCTTAT	2261
Qy	1861 ctgatgcacagaggagatctcgcaaaatbtggcggaacttgcagagcgtcttcaattcat	1920
Dp	2262 CTGATGCCAGGAGGATTCGGACAAATGGCGCACATGGCCGAGCGTCTTTCAATTTGAT	2321
Qy	1921 caccgatagagagagaaagaccgcccccttaaaacaagaatggttggagaatattctcat	1980
Dp	2322 CACCGATGAGAGAGAAAGACCGCCCCCTCAAAAACAAATAGTGTGAAGATATTCTGAT	2381
Qy	1981 cctggcttcagagagccttgcctttaaacaactgaa	2040
Dp	2382 CCTGGGCTTCAGAGACCTTGGCTTTACGGCGGAA	2441

[illegible]

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L. A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2			21583737	2 (bases 1 to 125020)	Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.	Direct Submission Submitted (05-Oct-2001) psychiatric Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA		location/Qualifiers 1..125020	1..125020

[illegible]



Db 50937 SSMWKSRRRRKSGAMGMKCYCTSSMRSTKSRSMKSSMWMGSKCYCYGWSCTKRS 50996  
Qy 880 accctcggagatattccctcgtatagcgcgagatcgaacccctggccattctcgac 939  
Db 50997 MGCYSTGSSMKMGWYCCGAGRSSRYCKSGSRMSMAGSGMSMGMTGMSGRG 51056  
Qy 940 cgtcaacgcatccatcctcgtagtaagatccgtcgtcgtcgtcccaagcgtccct 999  
Db 51057 CARGRMSATCSCRRKCKSKRSKSGCMRCSWYWKCMRSGRMKCKMSYTKCKGCSA 51116  
Qy 1000 agctcgcgaattctgtagatgagcggaacacttaccgtcccgatcgcgagc 1059  
Db 51117 TKSMAWYSCMRMRKGTKSSAMWYSCMCCSARAASGMYKCMKMGCGYGGKGSASTW 51176  
Qy 1060 ctctctgagatgccccggaggaagatgltcgttggaaacgagcgttaacagattgctctc 1119  
Db 51177 MYMCCSRACYYTTRKSSSRGWRKATRTMTTRSSSRMSWYKYAMMMRRMTTAgKKWKT 51236  
Qy 1120 aaacctgactgtagatgaatcaaccattat 1150  
Db 51237 KTTWSTKKKSKWGMYSRMMWMTTKYT 51267

RESULT 8  
AX024376 1969 bp DNA linear BCT 15-SEP-2000  
LOCUS Sequence 74 from Patent DE19846493.  
DEFINITION AX024376  
ACCESSION AX024376.1 GI:10184580  
VERSION  
KEYWORDS  
SOURCE Polyangium cellulosum.  
ORGANISM Polyangium cellulosum  
Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
Mycococcales; Sorangineae; Polyangiaceae; Polyangium.  
REFERENCE  
AUTHORS Beyer, S. and Mueller, R. J.  
JOURNAL Patent: DE 19846493-A 13-APR-2000;  
BIOTECHNOLOG FORSCHUNG GMBH (DE)  
FEATURES  
source 1..1969  
/organism="Polyangium cellulosum"  
/db\_xref="taxon:56"  
BASE COUNT 415 a 585 c 591 g 376 t 2 others  
ORIGIN

Query Match 1.8%; Score 46.2; DB 1; Length 1969;  
Best Local Similarity 50.7%; Pred. No. 0.23;  
Matches 111; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
Qy 520 gccgcagacggcgcttcgaccagaaaactgtctccttcgaggtgagagag 579  
Db 411 GCGCCGATCGCGCTCGCGAAGACGAGCTCGCCGCGCATCGCGGACGACATACC 470  
Qy 580 gtcatatgccaatcgccccgcgtcagcgcgtaacgcatcatatgaatgaac 639  
Db 471 GAGCAGAAACGATCGGACGACGAGCTCGAGCTCGAGCATATTTCATGACCCAC 530  
Qy 640 actcgaggcgcgcgatgctgtagatgtagcgaatggaatgagcgatcgac 699  
Db 531 CGCGCGGCGCGCGCGCGCGCATGACGACGACGATCAACGCGCTTGACCTTGAG 590  
Qy 700 gcaatgacggcatgaaaacacatgacggtgagagcgt 738  
Db 591 CCAACCGCGGAGTCCACACGAGAGAGCTGAGCTCTG 629

RESULT 9  
AX024269 1969 bp DNA linear PAT 15-SEP-2000  
LOCUS Sequence 74 from Patent DE19846493.  
DEFINITION AX024269  
ACCESSION

VERSION AX024269.1 GI:10184543  
KEYWORDS  
SOURCE Polyangium cellulosum.  
ORGANISM Polyangium cellulosum  
Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
Mycococcales; Sorangineae; Polyangiaceae; Polyangium.  
REFERENCE  
AUTHORS Beyer, S. and Mueller, R. J.  
JOURNAL Patent: DE 19846493-A 13-APR-2000;  
BIOTECHNOLOG FORSCHUNG GMBH (DE)  
FEATURES  
source 1..1969  
/organism="Polyangium cellulosum"  
/db\_xref="taxon:56"  
BASE COUNT 415 a 585 c 591 g 376 t 2 others  
ORIGIN

Query Match 1.8%; Score 46.2; DB 6; Length 1969;  
Best Local Similarity 50.7%; Pred. No. 0.23;  
Matches 111; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
Qy 520 gccgcagacggcgcttcgaccagaaaactgtctccttcgaggtgagagag 579  
Db 411 GCGCCGATCGCGCTCGCGAAGACGAGCTCGCCGCGCATCGCGGACGACATACC 470  
Qy 580 gtcatatgccaatcgccccgcgtcagcgcgtaacgcatcatatgaatgaac 639  
Db 471 GAGCAGAAACGATCGGACGACGAGCTGAGCTGAGCAGCATTTTCATGACCCAC 530  
Qy 640 actcgaggcgcgcgatgctgtagatgtagcgaatggaatgagcgatcgac 699  
Db 531 CGCGCGGCGCGCGCGCGCGCATGACGACGACGATCAACGCGCTTGACCTTGAG 590  
Qy 700 gcaatgacggcatgaaaacacatgacggtgagagcgt 738  
Db 591 CCAACCGCGGAGTCCACACGAGAGAGCTGAGCTCTG 629

RESULT 10  
I66494/c 7218 bp DNA linear PAT 28-DEC-1997  
LOCUS Sequence 14 from patent US 5670367.  
DEFINITION I66494  
ACCESSION I66494.1 GI:2724471  
VERSION  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.  
TITLE Recombinant fowlpox virus  
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
FEATURES  
source 1..7218  
/organism="unknown"  
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others  
ORIGIN

Query Match 1.8%; Score 45.8; DB 6; Length 7218;  
Best Local Similarity 3.8%; Pred. No. 0.38;  
Matches 14; Conservative 205; Mismatches 152; Indels 0; Gaps 0;  
Qy 1534 atagtgttaatgacttcgtgaacgagtgtgcaatataagcgttaagactgaataca 1593  
Db 1457 AAAGAGATAGAAGATTGTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398  
Qy 1594 cggctgacgagctgcaaaagcaagcgtgtgagcaacgcaacacacacacacacac 1653  
Db 1397 RRR 1338  
Qy 1654 aaagcgttaaatctcgttaaaaaagatgaaatcgtgggcatggtatagatcctcgag 1713



[illegible]

RESULT	11
LOCUS	CEP08G5
DEFINITION	Caenorhabditis elegans cosmid F08G5, complete sequence.
ACCESSION	32784 bp DNA linear INV 23-JAN-2007
VERSION	270682
KEYWORDS	Z70682.1 GI:1256464
SOURCE	HTG: Collagen; Transfer RNA; tRNA-Arg; tRNA-Gln; UDP-glucuronosyltransferase like.
ORGANISM	Caenorhabditis elegans. Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE	1 (sites)
AUTHORS	none.
TITLE	Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium
JOURNAL	Science 282 (5396), 2012-2018 (1998)
MEDLINE	99069613
REMARK	The <i>C. elegans</i> Sequencing Consortium.
REFERENCE	2 (bases 1 to 32784)
AUTHORS	Steward C.A.
TITLE	Direct Submission
JOURNAL	Submitted (06-APR-1996) Nematode Sequencing Project, Sanger

COMMENT

Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.

Current sequence finishing criteria for the *C. elegans* genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.

IMPORTANT: This sequence is not the entire insert of clone F08G5. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone F08G5 is at 1 in this sequence. The true right end of clone F08G5 is at 7448 in sequence Z70680.

The true left end of clone C25G4 is at 32680 in this sequence. The true right end of clone F28D1 is at 725 in this sequence. The start of this sequence (1..103) overlaps with the end of sequence Z70684 and of this sequence (32680..32784) overlaps with the start of sequence Z70680.

For a graphical representation of this sequence and its analysis see:- <http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=F08G5>

IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

FEATURES  
source

gene  
location/Qualifiers  
1. .32784  
/organism="Caenorhabditis elegans"  
/db\_xref="taxon:6239"  
/chromosome="IV"  
/clone="F0865"  
join(2838. .2935,2985. .3151,3200. .3582,3633. .3950,4009. .4101)  
/gene="F0865.1"  
join(2838. .2935,2985. .3151,3200. .3582,3633. .3950,4009. .4101)  
/gene="F0865.1"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4583.1"  
/db\_xref="GI:3875592"  
/db\_xref="SPRMBL:O192220"  
/translation="MEVNSRKLTIVHGEPIETVYCNMEVSRYPSELMKNTILFVYN  
NOAGPFLTRGDEDRNPLSTYRKFAVISEPVPDHLNRSTNTDWNVSVEKRGSS  
SQTSTSPQCDICDIDIGSVKKSAAARNLSHTYSTNIGSSRMPLATIANEFSGQVNPY  
NRPSASTVSSSTSGSAFVYLDDPSFGPSPNNSHLSGPCPSHSSSESGFSGSS  
SHSSQILSSFLNARNSPFDHSPPSANDIQIDQLNPVYAPQIKQADKCVQDQKQVD  
ERFPDPFLKRVKRVSDSKMARLQGMKRSIRKLNPDEDAFVKYKVDHQGTASS  
SSKIGPAIVQQTSDMEF"  
join(4263. .4329,4376. .4555,4855. .5113,5309. .5738,5788. .5943)  
/gene="F0865.2"  
join(4263. .4329,4376. .4555,4855. .5113,5309. .5738,5788. .5943)  
/gene="F0865.2"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4584.1"  
/db\_xref="GI:3875593"  
/db\_xref="SPRMBL:O19221"  
/translation="MEANDNKKRGWTLIDNVISCLLYVAGIYATFICPIVFLCEFY  
TAPILMLSTERKFKVKTAKRAGNLOFTVCRITGVYVMSGDLELTITCRALLP  
NLHGLGDFHITMTAADSGVNAAGRMFLVITYNMNLTSPGLMSSYSGNGLFVDSVPAP  
RAQTHLSIRHEDRIAYHEVDLRLCLYPSGRSLFLTKRNSFEKKRGTPEITPCAVP  
RGAAALSIKRVHGDPPNNPLKSNNGKPPLEYIVDTVLQTPDGNLPLUNETPNSAPN  
GGGCPAIIHVEIFKQDPMKDEELKNEFLFDYRKDKRLLESYVTSFSSEKSAOF  
PSMLIGHILITINGLEFALIVYFLFSFY"  
complement(join(6648. .6998,7613. .7761,7964. .8053,8357. .8473))  
/gene="F0865.3"  
complement(join(6648. .6998,7613. .7761,7964. .8053,8357. .8473))  
/gene="F0865.3"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4582.1"  
/db\_xref="GI:3875591"  
/db\_xref="SPRMBL:O19219"  
/translation="MSYQHEVYKATPEENDEIHAIIRGKPSFCKANIRVNOQFOER  
TYLDRKDLIRKMTSLSRKYVNELDIKRLDNEENNDPTOKLDEIKREKQIMRDEK  
HIIKGEISLTWNEDLRRRLAGEWKKIDADKSSAASEASSSMIGDDTPEISDSGQ  
KPYRETSM"  
complement(10702. .10773)  
/gene="F0865.t3"  
/note="CtG Gln Q-tRNA  
predicted using tRNAscan-se-1.11  
preliminary prediction  
similar to tRNA-Gln"  
/product="tRNA-Gln"  
complement(10702. .10773)  
/gene="F0865.t3"  
complement(11009. .11080)  
/note="t3"

gene  
location/Qualifiers  
1. .32784  
/organism="Caenorhabditis elegans"  
/db\_xref="taxon:6239"  
/chromosome="IV"  
/clone="F0865"  
join(2838. .2935,2985. .3151,3200. .3582,3633. .3950,4009. .4101)  
/gene="F0865.1"  
join(2838. .2935,2985. .3151,3200. .3582,3633. .3950,4009. .4101)  
/gene="F0865.1"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4583.1"  
/db\_xref="GI:3875592"  
/db\_xref="SPRMBL:O192220"  
/translation="MEVNSRKLTIVHGEPIETVYCNMEVSRYPSELMKNTILFVYN  
NOAGPFLTRGDEDRNPLSTYRKFAVISEPVPDHLNRSTNTDWNVSVEKRGSS  
SQTSTSPQCDICDIDIGSVKKSAAARNLSHTYSTNIGSSRMPLATIANEFSGQVNPY  
NRPSASTVSSSTSGSAFVYLDDPSFGPSPNNSHLSGPCPSHSSSESGFSGSS  
SHSSQILSSFLNARNSPFDHSPPSANDIQIDQLNPVYAPQIKQADKCVQDQKQVD  
ERFPDPFLKRVKRVSDSKMARLQGMKRSIRKLNPDEDAFVKYKVDHQGTASS  
SSKIGPAIVQQTSDMEF"  
join(4263. .4329,4376. .4555,4855. .5113,5309. .5738,5788. .5943)  
/gene="F0865.2"  
join(4263. .4329,4376. .4555,4855. .5113,5309. .5738,5788. .5943)  
/gene="F0865.2"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4584.1"  
/db\_xref="GI:3875593"  
/db\_xref="SPRMBL:O19221"  
/translation="MEANDNKKRGWTLIDNVISCLLYVAGIYATFICPIVFLCEFY  
TAPILMLSTERKFKVKTAKRAGNLOFTVCRITGVYVMSGDLELTITCRALLP  
NLHGLGDFHITMTAADSGVNAAGRMFLVITYNMNLTSPGLMSSYSGNGLFVDSVPAP  
RAQTHLSIRHEDRIAYHEVDLRLCLYPSGRSLFLTKRNSFEKKRGTPEITPCAVP  
RGAAALSIKRVHGDPPNNPLKSNNGKPPLEYIVDTVLQTPDGNLPLUNETPNSAPN  
GGGCPAIIHVEIFKQDPMKDEELKNEFLFDYRKDKRLLESYVTSFSSEKSAOF  
PSMLIGHILITINGLEFALIVYFLFSFY"  
complement(join(6648. .6998,7613. .7761,7964. .8053,8357. .8473))  
/gene="F0865.3"  
complement(join(6648. .6998,7613. .7761,7964. .8053,8357. .8473))  
/gene="F0865.3"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4582.1"  
/db\_xref="GI:3875591"  
/db\_xref="SPRMBL:O19219"  
/translation="MSYQHEVYKATPEENDEIHAIIRGKPSFCKANIRVNOQFOER  
TYLDRKDLIRKMTSLSRKYVNELDIKRLDNEENNDPTOKLDEIKREKQIMRDEK  
HIIKGEISLTWNEDLRRRLAGEWKKIDADKSSAASEASSSMIGDDTPEISDSGQ  
KPYRETSM"  
complement(10702. .10773)  
/gene="F0865.t3"  
/note="CtG Gln Q-tRNA  
predicted using tRNAscan-se-1.11  
preliminary prediction  
similar to tRNA-Gln"  
/product="tRNA-Gln"  
complement(10702. .10773)  
/gene="F0865.t3"  
complement(11009. .11080)  
/note="t3"

gene  
location/Qualifiers  
1. .32784  
/organism="Caenorhabditis elegans"  
/db\_xref="taxon:6239"  
/chromosome="IV"  
/clone="F0865"  
join(2838. .2935,2985. .3151,3200. .3582,3633. .3950,4009. .4101)  
/gene="F0865.1"  
join(2838. .2935,2985. .3151,3200. .3582,3633. .3950,4009. .4101)  
/gene="F0865.1"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4583.1"  
/db\_xref="GI:3875592"  
/db\_xref="SPRMBL:O192220"  
/translation="MEVNSRKLTIVHGEPIETVYCNMEVSRYPSELMKNTILFVYN  
NOAGPFLTRGDEDRNPLSTYRKFAVISEPVPDHLNRSTNTDWNVSVEKRGSS  
SQTSTSPQCDICDIDIGSVKKSAAARNLSHTYSTNIGSSRMPLATIANEFSGQVNPY  
NRPSASTVSSSTSGSAFVYLDDPSFGPSPNNSHLSGPCPSHSSSESGFSGSS  
SHSSQILSSFLNARNSPFDHSPPSANDIQIDQLNPVYAPQIKQADKCVQDQKQVD  
ERFPDPFLKRVKRVSDSKMARLQGMKRSIRKLNPDEDAFVKYKVDHQGTASS  
SSKIGPAIVQQTSDMEF"  
join(4263. .4329,4376. .4555,4855. .5113,5309. .5738,5788. .5943)  
/gene="F0865.2"  
join(4263. .4329,4376. .4555,4855. .5113,5309. .5738,5788. .5943)  
/gene="F0865.2"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4584.1"  
/db\_xref="GI:3875593"  
/db\_xref="SPRMBL:O19221"  
/translation="MEANDNKKRGWTLIDNVISCLLYVAGIYATFICPIVFLCEFY  
TAPILMLSTERKFKVKTAKRAGNLOFTVCRITGVYVMSGDLELTITCRALLP  
NLHGLGDFHITMTAADSGVNAAGRMFLVITYNMNLTSPGLMSSYSGNGLFVDSVPAP  
RAQTHLSIRHEDRIAYHEVDLRLCLYPSGRSLFLTKRNSFEKKRGTPEITPCAVP  
RGAAALSIKRVHGDPPNNPLKSNNGKPPLEYIVDTVLQTPDGNLPLUNETPNSAPN  
GGGCPAIIHVEIFKQDPMKDEELKNEFLFDYRKDKRLLESYVTSFSSEKSAOF  
PSMLIGHILITINGLEFALIVYFLFSFY"  
complement(join(6648. .6998,7613. .7761,7964. .8053,8357. .8473))  
/gene="F0865.3"  
complement(join(6648. .6998,7613. .7761,7964. .8053,8357. .8473))  
/gene="F0865.3"  
/note="predicted using GeneFinder"  
/codon\_start=1  
/protein\_id="CAAG4582.1"  
/db\_xref="GI:3875591"  
/db\_xref="SPRMBL:O19219"  
/translation="MSYQHEVYKATPEENDEIHAIIRGKPSFCKANIRVNOQFOER  
TYLDRKDLIRKMTSLSRKYVNELDIKRLDNEENNDPTOKLDEIKREKQIMRDEK  
HIIKGEISLTWNEDLRRRLAGEWKKIDADKSSAASEASSSMIGDDTPEISDSGQ  
KPYRETSM"  
complement(10702. .10773)  
/gene="F0865.t3"  
/note="CtG Gln Q-tRNA  
predicted using tRNAscan-se-1.11  
preliminary prediction  
similar to tRNA-Gln"  
/product="tRNA-Gln"  
complement(10702. .10773)  
/gene="F0865.t3"  
complement(11009. .11080)  
/note="t3"

gene  
location/Qualifiers  
1. .32784  
/organism="Caenorhab

```
/note="predicted using tRNAscan-SE 1.1.1  
preliminary prediction  
similar to tRNA-Gln"  
/product="tRNA-Gln"
```



[illegible]







Db 1970 GCTGCAGG 1977

RESULT 14  
AC008040/c

LOCUS 204917 bp DNA linear PRI 09-OCT-1999

DEFINITION Homo sapiens 3q26.2-27 BAC RPi11-379K17 (Roswell Park Cancer  
Institute Human BAC Library) complete sequence.

AC008040  
AC008040  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

HTG.

human.

Homo sapiens

Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 204917)  
Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,  
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,  
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,  
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,  
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,  
Forcum-Tansey,J., Frantz,P., Ganesha,R., Garcia,D.K., Gorrell,J.H.,  
Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,  
Hodgson,A., Hogue,M., Hollway,C., Hosak,H., Jackson,L.E.,  
Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondajewski,N., Kong,Y.,  
Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O.,  
Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G.,  
Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S., Nelson,A.,  
Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S.,  
Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M.,  
Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Simon,M.,  
Spark,A., Stamps,A., Sucgang,R., Taber,P., Taylor,T., Vasquez,L.,  
Vinson,R., Vo,O., Wahbah,M., Watlington,S., Weinstein,G.,  
Weinstock,I.R., Williamson,A., Worley,K., Wren,J., Wrenford,G.,  
Yu,M., Zhou,X., Naylor,S.L., Nelson,D. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished  
2 (bases 1 to 204917)  
Worley,K.C.  
Direct Submission  
Submitted (14-JUL-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

Worley,K.C.  
Direct Submission  
Submitted (24-SEP-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

Worley,K.C.  
Direct Submission  
Submitted (09-OCT-1999) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Sep 24, 1999 this sequence version replaced gi:5739555.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST

(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES  
Source Location/Qualifiers

1..204917

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="3q26.2-27"

/clone="RP11-379K17"

766..1132

/note="Region: od11d01.s1 Homo sapiens cDNA, AA810014"

2456..2779

/note="Region: zs55f12.s1 Homo sapiens cDNA, AA286777"

2554..2768

/standard\_name="G37704"

/db\_xref="dbSTS:55264"

6897..7126

/rpt\_family="MER20"

7488..8243

/note="Region: Unigene cluster containing AA400519 and  
AA165136"

11171..11471

/rpt\_family="AlusX"

/complement(13526..13835)

/rpt\_family="AlusG"

13556..15048

/note="Region: Unigene cluster containing A1683116,  
A1048825, AA278936, and W38986"

/complement(14286..14586)

/rpt\_family="AlusX"

15264..15890

/note="Region: Unigene cluster containing A1042019 and  
A1890489"

15800..15834

/rpt\_family="AT-rich"

16816..16837

/rpt\_family="AT-rich"

16911..17154

/rpt\_family="MIR"

/complement(17196..17318)

17598..17782

/rpt\_family="MER112"

18072..19357

/note="Region: Unigene cluster containing A1767291,  
A1522162, g311729 and AA354506"

18120..18226

/standard\_name="G24029"

/db\_xref="dbSTS:37867"

20464..20496

/rpt\_family="(CMAAAA)n"

20576..20882

/rpt\_family="Alu0"

/complement(20890..21061)

21061..21061

/rpt\_family="MER46A"



```

repeat_region      21460..21885
                    /rpt_family="L1PA13"
repeat_region      22470..22760
                    /rpt_family="Aluub"
repeat_region      24017..24179
                    /rpt_family="Charlielb"
repeat_region      24180..24635
                    /rpt_family="LTR22A"
repeat_region      24639..24744
                    /rpt_family="HERVK2"
repeat_region      24750..25603
                    /rpt_family="HERVK2"
repeat_region      25606..26058
                    /rpt_family="LTR22A"
repeat_region      26061..26174
                    /rpt_family="Charlielb"
repeat_region      26391..26559
                    /rpt_family="MIR"
repeat_region      26943
                    /rpt_family="MIR"
repeat_region      27836..28013
                    /rpt_family="MLT1I"
repeat_region      28215..28526
                    /rpt_family="Alusq"
repeat_region      28548..28687
                    /rpt_family="L1M4"
repeat_region      28688..28981
                    /rpt_family="Aluub"
repeat_region      28982..29281
                    /rpt_family="L1M4"
repeat_region      29282..29588
                    /rpt_family="L1M4"
repeat_region      29589..30042
                    /rpt_family="Alusx"
repeat_region      30055..30213
                    /rpt_family="L1M4"
repeat_region      30213
                    /rpt_family="L1M4"
repeat_region      30264..30568
                    /rpt_family="Aluub"
repeat_region      30587..30601
                    /rpt_family="MLT1C"
repeat_region      31022
                    /rpt_family="MSTB"
repeat_region      31023..31094
                    /rpt_family="MLT1C"
repeat_region      31096..31402
                    /rpt_family="Alusx"
repeat_region      31404..31594
                    /rpt_family="L1M2"
repeat_region      31605
                    /standard_name="D3S3468"
                    /db_xref="dbSTS:52020"
                    /complement(31595..31673)
repeat_region      31673
                    /rpt_family="MSTB"
repeat_region      31674..31959
                    /rpt_family="Alusx"
repeat_region      31960..32324
                    /rpt_family="MSTB"
repeat_region      32325..32567
                    /rpt_family="L1M2"

Query Match      1.7%; Score 43; DB 9; Length 204917;
Best Local Similarity 56.8%; Pred. No. 4.6;
Matches 79; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

Db 143098 TTACTGTCTATTATGCC 143080

```

RESULT 15
AE005932      13291 bp DNA linear BCT 28-MAR-2001
LOCUS      Caulobacter crescentus section 258 of 359 of the complete genome.
DEFINITION
AE005932 AE005673
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
    Caulobacter crescentus.
    Caulobacter crescentus
    Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
    Caulobacter.
REFERENCE
    1 (bases 1 to 13291)
    Niemman,W.C., Feldhlyum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E.,
    Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R.,
    Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D.,
    Ely,B., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L.,
    Ely,D.H., Kolonay,J.F., Smit,J., Graven,M., Khouiri,H., Shetty,J.,
    Berry,K., Utterback,T., Tran,K., Wolf,A., Vamathevan,J.,
    Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and
    Fraser,C.M.
    Complete genome sequence of Caulobacter crescentus
    Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
2113698
TITLE
    Complete genome sequence of Caulobacter crescentus
JOURNAL
    Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
REFERENCE
    2 (bases 1 to 13291)
    Niemman,W.C., Feldhlyum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J.,
    Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I.,
    Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B.,
    Laub,M.T., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L.,
    Laub,M.T., Kolonay,J.F., Smit,J., Graven,M., Khouiri,H., Shetty,J.,
    Berry,K., Utterback,T., Tran,K., Wolf,A., Vamathevan,J.,
    Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and
    Fraser,C.M.
    Direct Submission
    Submitted (31-JAN-2001) The Institute for Genomic Research, 9712
    Medical Center Dr, Rockville, MD 20850, USA
FEATURES
    source
        1..13291
            Location/Qualifiers
                /organism="Caulobacter crescentus"
                /db_xref="taxon:69394"
                79..3009
                /gene="CC2638"
                79..3008
                /note="Identified by match to PRM protein family HMM
                PF00675"
                /codon_start=1
                /transl_table=11
                /product="peptidase, M16 family"
                /protein_id="AAK24605.1"
                /db_xref="GI:13424217"
                /translation="MISASRLAIVAAAGISVAACSQLPKSPILKKDAGASAPKTP
                DAPALGPKVTDLKPQWAOELSDVAPDPMRFGVLPNMRKRLAKKNATPPQOARLT
                MIDAGSMMEDDOGLAHLFEMHAFNGSNVPEGEMIKLEHGLAFGADTNAISTPD
                ETIYOLDLPKTDIDTDTSLMLRFAAGELTIAPEVADREKGVLSSEETRVTPYRV
                AIRTLSONWEGOLPKPRKIPRGTEVTKPAQIRPSEYRPERVLAAGDDEVD
                AMEAKIKGKGDWYKGRPNKDPDVGPAKRGRTAFMEAGAPMSIQMTWTRKEGL
                LETKAVDERDTLENLGFVILNRLQAVGSAPEPFTAGAGFAGDGOAGAVRTPGATA
                QGRWREALTAALDAEORAIÖYGRDDELREIASIRACILVAAAGEAORPISLANO
                LGVTLGDGEVATSPSONIAAFDFVKGFLAERNAVALKSAFVSGCLDVALVLRPNG
                LEPAILKAVEELKQPAAPPIAGCVIWPVSSSGPFGKAEQDEVDLAVLRPNG
                VRLVTKPKTKGVEDDOVIVKVRAGHLIDMPEDKOSPMSGSFLEGLKQITQDMKRY
                LTGKINNAOLGVEDDATTNGRTRPELSTELQVLAFAATEPMPREANRRTKISYGT
                LHDQLOSTGCVLRDGLGIMHGSDMRFPREQLASASLIDLKAAVNPPLAKGDLQ
                VLVIGDVTVDALIAAVADTIGALPARPCDPLGAKKAPAPARSKBPVIRTHRGPDQ
                AALPMWMTDLDLFSNLRSDVSLGVGYMLRLTDELREKQATSPVMTATASVVEF
                DMGVLAVSLVPPPEKLDGVVASKIADIRKDVSADELDRAKKRIRDAIEKARVTN
                EYWGVALSGAHTDRLDLATRSVYAGLSRVTPADVQAAQTYLADEKSWLLKVPMAA
                AAAA"
                complement(3071..4000)

```

gene



gene  
CDS  
/gene="CC2639"  
complement(3071..4000)  
/gene="CC2639"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK24606.1"  
/db\_xref="GI:13424218"  
/translation="MKAPKLMTPAWTLTALGMLGGLLIGCAAYAGCTKACPPPPCCKPPEPPKPPAPPTGCCGGHHNNVILPCVVNISTSYIVANLTAAPIANSGATSGAAVCASTGAPPVVYVGGGSGCYSGSPATGVITGLNNEGAVYRRQKVEATRTPTRVIVRAICLDRDDPPHASOVTIPDREIDENIDELRCTIAGARLVINEYGGHIDITSGVAGQFFIEIKRGALVYARGAAGAIVSCROGLPARNCENSLIRREGAGVKILMILTETVYVAREEYEOALSTSEFSLDGGVGVVY"  
3970..4722  
/gene="CC2640"  
3970..4722  
/gene="CC2640"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK24607.1"  
/db\_xref="GI:13424219"  
/translation="MRACSVSPASSGVSLHLSYIRLRSKNRAEAAAPWICARMGRSSSWPSPSPKPRVAVENSRARTRQGTMMHPSSTERLIDYRAKCGOALPRRADIDPGFLELLPOVYLVGREGKLEPFLAGFVSDLHARDLRGHDALSLMALHRLKSLADVAARRRTPVVTVDIRAHGVSPVSEVLEFAPLAGASGETDRFLGLQPIAMARILMGRRAYELGLREIKPLGDSQNDMPRLRLATLDGRRIA"  
complement(4719..5663)  
/gene="CC2641"  
complement(4719..5663)  
/gene="CC2641"  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="oxidoreductase, Aldo/Keto reductase family"  
/protein\_id="AAK24608.1"  
/db\_xref="GI:13424220"  
/translation="MRHRPLRGSLTLPRLFGVGVGMTADEMSHRLDAPIDGCFNAVDTADVYSAMVPGHAGESPAVIGRWLKASGRKDSVLTVTKYAMMKRPGLSAANITEAVALSESLRGLDTQDIDLYQSHQDDAETPODETLSAEDRLVQAGKVAVGSNTPARLKASIDVSTASGLARVETITQPFNLVNDQVGEALAEALANEGGIIPYVSLAGFLTLAGYREGSDIEGRARGRTILRDVMDNGSAVLAALDEAAALAAVAGSPAVALAMIMRHBSITAPYASPTVYQDLDLMGAARLDLPDDIMATLDAAGR"  
complement(5674..7341)  
/gene="CC2642"  
complement(5674..7341)  
/gene="CC2642"  
/note="identified by match to PFAM protein family HMM PF00732"  
/codon\_start=1  
/transl\_table=11  
/product="choline dehydrogenase"  
/protein\_id="AAK24609.1"  
/db\_xref="GI:13424221"  
/translation="MASRPRDVIYVAGSACCVLAARLTEDPNIKVILLFEXGKNKSTILYKMPAGVQGLIKDKGEONKCFWTEAEPHLNDRKIMPRFGCLGSSALINMITYIRSHADYDQMRKHTLGMVSYEVLPERKSTETHAGDAVHGSGPLHVSAGESDSFPYSTLIVEADQRQMRKHTKDFNGVYQOGEFQYDLTRDQRMSSAAAYILNOALSRLNLCVTEETACTPTIILDKRAVGEVYVVGSKSREKQYAVDAEVLISAGAVSQPOLIOLSGIAAEDDIALPHGIIAANHEKSGVANLOHLDVYCVMTAKNLKTAVSANKGLNKLGIVGNVMPFGKRGIGROOLFESGAEKRSRPLDRDRLQITGVYALIMDHKKVYVEKDGFTLHNCQLRPSSRKVGILRSADPPDPTIIGNTLATEDEBRALIREGVLAKRETVAAQADPTIDATYAVGADVKSDDADIDAWIRSKAETIYHPVGTGCRMGVAGPMVAVVDQLRVQGVQRLVDAVAGMPTLIGNTNAETPIIMASRAADLIRGKTPAPRLDVYVEDGRAVAAE"  
complement(7633..8163)  
/gene="CC2643"  
complement(7633..8163)  
/gene="CC2643"  
/note="identified by match to protein family HMM"  
/codon\_start=1

					/transl_table=1 /product="Bacterial transferase family protein" /protein_id="AAK24610.1" /db_xref="GI:13424222"
					/translation="MTVYSLGATPTPLPAQGEYWIAPASVMGNVLTKNNAISIMWGANVARGNDPIITGENSNVDGSYLHDDIDCAPLITIGNNNYIGHMVMHGCTIGDGLIGSYLVNGAKTIGKNCLICAGALTTECKETIPDNMSVMGAFKKYVERIGEONAHMLAASALHYVENMKRYVRDLKTIVE"
gene					8525..9277 /gene="CC2644"
CDS					8525..9277 /gene="CC2644" /note="similar to GB:D10391.SP:PJ3544.PID:285774.GB:U00096.PID:I651507; identified by sequence similarity; putative"
					/codon_start=1 /transl_table=1 /product="PhoP-related protein" /protein_id="AAK24611.1" /db_xref="GI:13424223"
					/translation="MARBEAHEIQYDYDAKYRRLPVPHRKSGSPLOGAASGDSRDOSYLKTIKRSPGOALMEAIIDEKENLYMNLGPGRGTGKTIKRALEADKRVSTIVLSRPAAVEAGEISQILPGDEMDKLAPLRPLYALDIRGLTSVAIVLAAMAGAEIAIAYWGPRGRLLNNAAVFIDEAQNCYTQMOLKMILLTRLMSTINVTVGDPNSDLLPISGLAEVAAAKFOAVPNIAVVARLDNRDIYRHPLVAEMLGVL"
gene					9453..9758 /gene="CC2645"
CDS					9453..9758 /gene="CC2645" /note="identified by Glimmer2; putative"
					/codon_start=1 /transl_table=1 /product="hypothetical protein" /protein_id="AAK24612.1" /db_xref="GI:13424224"
					/translation="MSLIKSALFATTTLLALAASPVSASEQFVPPOSAGTMSOLIDMGRQLTVTAGARATKVOMERAVRAALINSRCXDAYLLAASEQDERLATRYEEVCAPR"
gene					9853..11853 /gene="CC2646"
CDS					9853..11853 /gene="CC2646" /note="identified by match to TIGR protein family HMW TIGR00733"
					/codon_start=1 /transl_table=1 /product="Oligopeptide transporter, OPT family" /protein_id="AAK24613.1" /db_xref="GI:13424225"
					/translation="MLMADTPAPRETLIRGVLLGITLTTFYAQVYVGLKVLGPFA TAIPAAVISMALRAPFTSTIOENNVIOTIASAGTSAVFVPLGLIMGMAWNPFLPFGACAVGGIIIGVNTYPIELRALYNQSNLPREGVAAAPVLKVGSGRGAEGKRGGLAIIIGSALSALFPALGALKAKLEPAEATIAEFKRLGASGATSIGSSLSIALMGACHLMGTIVGAMFTGLFIAMAILVPLITLVTPPEADAATHALTWKSOYRFGLAGVTAIALMTIKLVGPRIITSGLSAFAPAAQAARKAGAKLPREODIPIGIVLSVLLAPAGAWFELAHFPIITGCP IASLTTPVLAIVIGIVLFAGLAAVCGVAGVIGSSNSPSGIALITSVLASPLTWGMVGRGVIGDPVTKALIFALVVTTCVLAIVAVANADIODLKTOGLDVLA TPXKUGCNLDMLGILGYCAIVTGILVANADAILRKSHERTSLPPRAGVGLATIVPSYSYTA PVVVGALAGWLTERKVVSKDRAPAESRRGLVLAGFTGESLFWNSIALLLTVSGKDD PLALPFAPSEHVHGFSLIAAAAIIVVGFYRMARRAKAGAKAMEA"
gene					11860..12507
Query Match	1.7%	Score 42.8:	DB 1;	Length 13291;	
Best Local Similarity	52.2%;	Pred. No. 3;			
Matches	95;	Conservative	0;	Mismatches	87;
				Indels	0;
				Gaps	0;
QY	647	gcggccgacatgtcctgatgtcgcccatgaattcggtctggccgacacatcgcgcaatcg	706		
	I I I I I	I I I I I	I I I I I	I I I I I	I I I I I
Db	11408	GGGGCCTGTGGGCATATGGTGCCTGAGTCCGGCCGTGGGCTGTGCCCGTCGACGCCGATCC	11467		
QY	707	accgcatgaaaaaccatgacgctgaagactgaaaccgcgctgttcgaggactagccgctg	766		
	I I I I I	I I I I I	I I I I I	I I I I I	I I I I I
Db	11468	TCCGCAGACGCCACGACGCCCTTCACGCTCCGCCCTGGGCTGGGCTGGGCGATCT	11527		







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 19:55:13 ; Search time 279.42 Seconds  
(without alignments)  
15465.871 Million cell updates/sec

Title: US-09-826-206-4  
Perfect score: 2517  
Sequence: 1 ggcacatgggtcgaattcat.....tgaggaaataagggaacgtg 2517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues  
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
N.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2517	100.0	2517	24	AA518309
2	2517	100.0	8509	22	AA167035
3	2517	100.0	8509	24	AA518307
4	2007	79.7	2112	24	AA518306
5	2007	79.7	5859	24	AA518308
6	40.4	1.6	731	13	AAO21036
7	40.4	1.6	1980	23	AA525212
8	40.4	1.6	2008	23	AA581796
9	39.8	1.6	10886	24	ABL34134

10	39	1.5	6050	24	ABL34011	Human immune syste
C 11	38.8	1.5	4590	22	AAH24065	Yeast AOD9604-asso
C 12	38.2	1.5	2047	23	AA587283	DNA encoding novel
C 13	37.8	1.5	513445	22	AA161373	Soybean 318013 reg
C 14	37.6	1.5	4043	22	AA526561	DNA encoding human
C 15	37.6	1.5	4044	22	AA564182	Human secreted pro
C 16	37.6	1.5	4049	21	AA559064	Human secreted pro
C 17	37.4	1.5	1456	24	AA517123	Replicon sequence
18	37.4	1.5	1458	24	AA517226	Replicon sequence
19	37.4	1.5	7029	24	AA517119	Ketogulonigenium e
20	36.8	1.5	858	23	AA554178	Pseudomonas aerugi
21	36.6	1.5	6216	19	AAV06946	Canine herpes viru
C 22	36.6	1.5	14000	21	AAA09020	A. thaliana ATPAC
C 23	36.6	1.5	36471	21	AAA81453	N. meningitidis pa
C 24	36.6	1.5	349980	21	AA521611	Neisseria meningit
C 25	36.4	1.4	2035	23	ABL10875	Drosophila melanog
C 26	36.4	1.4	4146	23	ABL10874	Drosophila melanog
C 27	36.2	1.4	1310	21	AA548039	Zea mays DNA fragm
C 28	35.8	1.4	4920	24	ABL34285	Human immune syste
C 29	35.8	1.4	4403765	22	AA196683	Mycobacterium tube
C 30	35.8	1.4	441529	22	AA196682	Mycobacterium tube
C 31	35.6	1.4	6692	23	ABL19806	Drosophila melanog
C 32	35.4	1.4	857	17	AA534530	Human Fas soluble
C 33	35.4	1.4	920	17	AA534529	Human Fas soluble
C 34	35.4	1.4	926	21	AA534528	Human ORF3103
C 35	35.4	1.4	975	17	AA534527	Human Fas soluble
C 36	35.4	1.4	1104	17	AA534527	Human Fas soluble
C 37	35.4	1.4	1167	17	AA534526	Human Fas antigen
C 38	35.4	1.4	1457	21	AA539167	Apoptobody3sc fusi
C 39	35.4	1.4	1982	22	AAH15840	Human CDNA sequenc
C 40	35.4	1.4	2471	16	AAO93879	Fas-delta-TM CDNA.
C 41	35.4	1.4	2471	20	AA524878	Soluble Fas recept
C 42	35.4	1.4	2534	13	AAO29597	Human cell surface
C 43	35.4	1.4	2534	16	AAO95297	Plasmid p58 contg
C 44	35.4	1.4	2534	17	AA516303	hfas coding sequen
C 45	35.4	1.4	2534	18	AAV07002	Human Fas antigen

ALIGNMENTS

RESULT 1  
AA518309 standard; DNA; 2517 BP.  
XX  
AC AA518309:  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE DNA region of Ketogulonigenium plasmid pADM291 supporting replication.  
XX  
KW Cloning vector; Ketogulonigenium replicon; endogenous plasmid; pADM291;  
KW Transformed host cell; Escherichia coli; plasmid vector replication; ds.  
XX  
OS Ketogulonigenium sp. strain ADM291-19.  
XX  
PN WO200177347-A2.  
PD 18-OCT-2001.  
XX  
PF 05-APR-2001; 2001WO-US11059.  
XX  
PR 05-APR-2000; 2000US-194625P.  
XX  
PA (ARCH ) ARCHER-DANIELS MIDLAND CO.  
XX (DELI/) D'ELIA J.  
XX D'Elia J;  
XX  
DR WPI; 2002-049150/06.  
XX  
PT Novel nucleic acid vector comprising Ketogulonigenium replicon found on a specific deposited endogenous plasmid, useful for producing



PT polypeptides and/or transcripts by culturing host cells transformed  
PT with vector -

XX  
PS  
Claim 4: Fig 4; 66pp; English.

XX The present invention relates to the isolation of vectors comprising  
CC a Ketoglutaenium replicon found on the endogenous plasmid, pADM291.  
CC The invention also describes methods of transforming host cells with  
CC the vectors and producing polypeptides and/or antisense transcripts by  
CC culturing the transformed host cells. The vectors are useful for  
CC transforming a host cell by conjugation or electroporation.  
CC The vectors which have a replicon functional in both Ketoglutaenium  
CC and Escherichia coli, enable the cloning of certain genes of  
CC Ketoglutaenium in E. coli as the latter is an efficient host for  
CC amplification of vector DNA. The present DNA sequence represents the  
CC region of Ketoglutaenium endogenous plasmid pADM291 that supports  
CC plasmid vector replication.

XX Sequence 2517 BP: 657 A; 589 C; 683 G; 588 T; 0 other;  
SO

100 08: Score 2517; DB 24; Length 2517;

Query match	100.00%	Pred. No. 0;		
Best Local Similarity	100.00%			
Mismatches	0;	Indels	0;	Gaps

QY	1	ggcaatbagggtcgaatatcatagaaattctgtgtgaagtgctgagcggctctgacaggggtg	60
Db	1	ggcaatbagggtcgaatatcatagaaattctgtgtgaagtgctgagcggctctgacagaggggtg	60
QY	61	ctgcgcggagatctctggtcttcaggttagggcgcgaatgagaggttttagttccccctg	120
Db	61	ctgcgcggagatctctggtcttcaggttagggcgcgaatgagaggttttagttccccctg	120
QY	121	tatcgccctcgcgtgcgcatttgggtgctacatctgcgcgacatatgatattccgtcagag	180
Db	121	tatcgccctcgcgtgcgcatttgggtgctacatctgcgcgacatatgatattccgtcagag	180
QY	181	gattactcgatagttcttcgcgcgtgcgggtctgttcgggtctgtgcggcttgcgcgctgtgc	240
Db	181	gattactcgatagttcttcgcgcgtgcgggtctgttcgggtctgtgcggcttgcgcgctgtgc	240
QY	241	gggcctgcctctctgtctccgcgcgtgctctcaattttcaacaatcaaaaaatggcgaagc	300
Db	241	gggcctgcctctctgtctccgcgcgtgctctcaattttcaacaatcaaaaaatggcgaagc	300
QY	301	ccctctgttctatagttcttatagtttcaatagcaaaattacacataattacataact	360
Db	301	ccctctgttctatagttcttatagtttcaatagcaaaattacacataattacataact	360
QY	361	attcgcttaaaaggaggagtaatttggcgcgcgaaaaaggagtaatttggcgcgcaaaaaggagt	420
Db	361	attcgcttaaaaggaggagtaatttggcgcgcgaaaaaggagtaatttggcgcgcaaaaaggagt	420
QY	421	aattggccgcgaaaaggagtaatttgggcgcgtatcgttttfttaccatgvgggagaaatcc	480
Db	421	aattggccgcgaaaaggagtaatttgggcgcgtatcgttttfttaccatgvgggagaaatcc	480
QY	481	cccttaactatcttctcccatctgggaaagacaacaagtgccgcgacacccggctctgcac	540
Db	481	cccttaactatcttctcccatctgggaaagacaacaagtgccgcgacacccggctctgcac	540
QY	541	cagacaaaaactgtgctccctgcgcggaggttggcgaagggtctatatgtgcacatccgcc	600
Db	541	cagacaaaaactgtgctccctgcgcggaggttggcgaagggtctatatgtgcacatccgcc	600
QY	601	cgcctggaagcgtctcaagctcatgcatcttataatgatagccacttgcgcggtgcgcgcatgtgct	660
Db	601	cgcctggaagcgtctcaagctcatgcatcttataatgatagccacttgcgcggtgcgcgcatgtgct	660
QY	661	gatgatgtgcgcgaatgaatgatgcgttgcgcgaattctgcgcgaatctgacatgaaac	720
Db	661	gatgatgtgcgcgaatgaatgatgcgttgcgcgaattctgcgcgaatctgacatgaaac	720

[illegible]



Db	1801	aaagacccccaagaattcagcgcattcagacgcgtgaaattccttgagccttaccgattccttat	1860
Oy	1861	ctgtatgcaggaggaatcgcgacaatctgcgcgacatggtccagatccttccaatttcat	1920
Db	1861	ctgtatgcaggaggaatcgcgacaatctgcgcgacatggtccagatccttccaatttcat	1920
Oy	1921	caccgatgatgagaggaagaccgccccctccaataaagaatctgtgtgagatlatctgat	1980
Db	1921	caccgatgatgagaggaagaccgccccctccaataaagaatctgtgtgagatlatctgat	1980
Oy	1981	ccctgtgcttcaaggagccttgctttaaacctggaatcagcattccagatctgtataag	2040
Db	1981	ccctgtgcttcaaggagccttgctttaaacctggaatcagcattccagatctgtataag	2040
Oy	2041	aagtaaatatagccacaatagagcggccatttccattacaataagctccatctgat	2100
Db	2041	aagtaaatatagccacaatagagcggccatttccattacaataagctccatctgat	2100
Oy	2101	caatataaaglatatgatattcatcaaatctggaagaattatcatctgatacagagatcatca	2160
Db	2101	caatataaaglatatgatattcatcaaatctggaagaattatcatctgatacagagatcatca	2160
Oy	2161	cagcatctgttttctgtattcttaagtgtcaataactatgcgtgtgcccctaagaagatt	2220
Db	2161	cagcatctgttttctgtattcttaagtgtcaataactatgcgtgtgcccctaagaagatt	2220
Oy	2221	gtactattgcagatctccacctgggtttccagacgagatattacagctttgtgaagcgg	2280
Db	2221	gtactattgcagatctccacctgggtttccagacgagatattacagctttgtgaagcgg	2280
Oy	2281	gtctttctgcagagagagacagcttttctgattgtttaaataacctgcagatctcatg	2340
Db	2281	gtctttctgcagagagagacagcttttctgattgtttaaataacctgcagatctcatg	2340
Oy	2341	atgtttaaataacacagacatcatctcgggtggacaacacaccttctatgtgaagatgatattg	2400
Db	2341	atgtttaaataacacagacatcatctcgggtggacaacacaccttctatgtgaagatgatattg	2400
Oy	2401	gccacagaagctactgtgcgaaaaacgcgatatctgagtcacagagcccgcactttagagcgg	2460
Db	2401	gccacagaagctactgtgcgaaaaacgcgatatctgagtcacagagcccgcactttagagcgg	2460
Oy	2461	aagcctataacgagttacccaacacatcagaagccagatctgaggaataatgggaacgtg	2517
Db	2461	aagcctataacgagttacccaacacatcagaagccagatctgaggaataatgggaacgtg	2517
RESULT 2			
AAI67035			
ID	AAI67035 standard; DNA: 8509 BP.		
XX	AAI67035;		
AC	11-FEB-2002 (first entry)		
XX	DE Nucleotide sequence of an endogenous ketogulonigenium plasmid.		
XX	Ketogulonigenium; plasmid, strain NRRL B-30035; vector; ds.		
OS	Synthetic.		
XX	W0200177159-A2.		
PN	18-OCT-2001.		
XX	05-APR-2001; 2001WO-US11058.		
XX	05-APR-2000; 2000US-194624P.		
XX	(UNMS ) UNIV MICHIGAN STATE.		
XX	(ARCH ) ARCHER-DANIELS MIDLAND CO.		

PI	Schmidt TM, Stoddard SF;
XX	
DR	WPI: 2001-657165/75.
XX	
PT	New nucleic acid comprising the sequence of a Ketogulonigenium plasmid
PT	designated pKUM291 is endogenous to microorganism strain NRRL B-30035
PT	
XX	
PS	Claim 1; Fig. 1A-E; 14pp; English.
XX	
CC	The present sequence represents a purified nucleic acid of a
CC	Ketogulonigenium plasmid endogenous to strain NRRL B-30035. The
CC	polynucleotide can be used in a vector. The vector comprises either
CC	(i) the polynucleotide and a marker gene or (ii) the polynucleotide, a
CC	promoter, a transcription terminator and a discrete series of restriction
CC	endonuclease recognition sites between the promoter and transcription
CC	terminator.
SQ	Sequence 8509 BP; 2065 A; 2239 C; 2345 G; 1860 T; 0 other:

[illegible]



QY 721 catgacggtgagagccttgaccccgctgttcgaggaagtacgacgctgagctgttgacccat 780  
|||||  
Db 3675 catgaccgttgagagccttgaccccgctgttcgaggaagtacgacgctgagctgttgacccat 3734  
QY 781 gatgacccctcaagaatgatcgtgacagctcgagcttgcgtatgagcgcgcaatagac 840  
|||||  
Db 3735 gatgacccctcaagaatgatcgtgacagctcgagcttgcgtatgagcgcgcaatagac 3794  
QY 841 taaccgcaaggagcaagcgagcgaactcctctagctgtgagccttcggggatcatcctgt 900  
|||||  
Db 3795 taaccgcaaggagcaagcgagcgaactcctctagctgtgagccttcggggatcatcctgt 3854  
QY 901 cgtatgacgagcgagctcgaaacacactgggacatctcgacccgtcaaacggtatccatctc 960  
|||||  
Db 3855 cgtatgacgagcgagctcgaaacacactgggacatctcgacccgtcaaacggtatccatctc 3914  
QY 961 ggtagtaagtaattccgtgtcgtcgttccagcaagctctctagtctcgcaacttgatcag 1020  
|||||  
Db 3915 ggtagtaagtaattccgtgtcgtcgttccagcaagctctctagtctcgcaacttgatcag 3974  
QY 1021 atgagcgcgaaaacaccttaacgctcccgagcttgcgagcgctcctcttgagtgcccgagga 1080  
|||||  
Db 3975 atgagcgcgaaaacaccttaacgctcccgagcttgcgagcgctcctcttgagtgcccgagga 4034  
QY 1081 aagatggtctgttgaaacgacgttaacagatctgtctcaaaccttgcaactggatgagatc 1140  
|||||  
Db 4035 aagatggtctgttgaaacgacgttaacagatctgtctcaaaccttgcaactggatgagatc 4094  
QY 1141 aaccttatacgtctctgacattgacggcaagcgacacagagatttgccgtaagctgtgca 1200  
|||||  
Db 4095 aaccttatacgtctctgacattgacggcaagcgacacagagatttgccgtaagctgtgca 4154  
QY 1201 agctgacataagctctgggaagtgaaaaagcaacccacgctccgacagcgcgagctgagc 1260  
|||||  
Db 4155 agctgacataagctctgggaagtgaaaaagcaacccacgctccgacagcgcgagctgagc 4214  
QY 1261 ggttccaaggtcgtcgcgagatgtcgttcgacagagggcgagcggaacagatagcccccctc 1320  
|||||  
Db 4215 ggttccaaggtcgtcgcgagatgtcgttcgacagagggcgagcggaacagatagcccccctc 4274  
QY 1321 ttccccaagcgcgcgaggtacacatacagttccacagttcgagctgaagctgaacgctctct 1380  
|||||  
Db 4275 ttccccaagcgcgcgaggtacacatacagttccacagttcgagctgaagctgaacgctctct 4334  
QY 1381 ggcagcaacaagagacaacagatctgacgtcagactcccgcgcttctgctggggagaga 1440  
|||||  
Db 4335 ggcagcaacaagagacaacagatctgacgtcagactcccgcgcttctgctggggagaga 4394  
QY 1441 ggcgtgcgtctgagacgtcgaacaacatcgaaacatgcttttagattctctgcgaagta 1500  
|||||  
Db 4395 ggcgtgcgtctgagacgtcgaacaacatcgaaacatgcttttagattctctgcgaagta 4454  
QY 1501 ggggaaggtttgagttttgaggtatttcaacgcgcaatagtgttaatgacttctgtgaacg 1560  
|||||  
Db 4455 ggggaaggtttgagttttgaggtatttcaacgcgcaatagtgttaatgacttctgtgaacg 4514  
QY 1561 atgtgcaataatagcggtgaagaactaataatacagcgtgacagcggtgcgaagcaacg 1620  
|||||  
Db 4515 atgtgcaataatagcggtgaagaactaataatacagcgtgacagcggtgcgaagcaacg 4574  
QY 1621 gtgtgacacgcgcaacatcatcactcggtcgcttaaaaagcggtlaaaattccggttaaaaag 1680  
|||||  
Db 4575 gtgtgacacgcgcaacatcatcactcggtcgcttaaaaagcggtlaaaattccggttaaaaag 4634  
QY 1681 atgaactgtgggcatgtgttatagatctctgcaaatgtgcacagagtgcttcccccattt 1740  
|||||  
Db 4635 atgaactgtgggcatgtgttatagatctctgcaaatgtgcacagagtgcttcccccattt 4694  
QY 1741 caaagaaatacaccgaaacacttaacacgcaagtataatggtaagcgtgatgaacacatg 1800  
|||||  
Db 4695 caaagaaatacaccgaaacacttaacacgcaagtataatggtaagcgtgatgaacacatg 4754  
QY 1801 aatgacctcagaataatcagcgcatagagcgtgaagttcgactttagcgagatgcttat 1860  
|||||

Db 4755 aatgacctcagaataatcagcgcatagagcgtgaagttcgactttagcgagatgcttat 4814  
|||||  
QY 1861 ctgattgccaaggagatctgcgacaataatgacgacatgacgacgtgcttcaatttcat 1920  
|||||  
Db 4815 ctgattgccaaggagatctgcgacaataatgacgacatgacgacgtgcttcaatttcat 4874  
QY 1921 caacgataagagaggaagaaacgccccctcaaaaaaagaatgtgtgaagatattctgat 1980  
|||||  
Db 4875 caacgataagagaggaagaaacgccccctcaaaaaaagaatgtgtgaagatattctgat 4934  
QY 1981 cctgggtctcaagagccttgccctttaaacctgaatacgaatctttagcgtcgtataag 2040  
|||||  
Db 4935 cctgggtctcaagagccttgccctttaaacctgaatacgaatctttagcgtcgtataag 4994  
QY 2041 aagtaataatagccacaataagcgcgcatcttccattccatacagctcatcgtat 2100  
|||||  
Db 4995 aagtaataatagccacaataagcgcgcatcttccattccatacagctcatcgtat 5054  
QY 2101 caatatcaagtaattgatattcaatgagagaagaatttaacatgatacagagatcatca 2160  
|||||  
Db 5055 caatatcaagtaattgatattcaatgagagaagaatttaacatgatacagagatcatca 5114  
QY 2161 cagcattgttttgatttattcttaagtgtcaacataactatgctgcccctaaagaagatt 2220  
|||||  
Db 5115 cagcattgttttgatttattcttaagtgtcaacataactatgctgcccctaaagaagatt 5174  
QY 2221 gtactatgacgatatccaccttggtttcagacgcaataataacagcttgcgaagcgg 2280  
|||||  
Db 5175 gtactatgacgatatccaccttggtttcagacgcaataataacagcttgcgaagcgg 5234  
QY 2281 gttttttgccaagagagagacagcttttctgagtggttaataacgctgacgatatctatg 2340  
|||||  
Db 5235 gttttttgccaagagagagacagcttttctgagtggttaataacgctgacgatatctatg 5294  
QY 2341 atgtgtaacatacagacgatatccggggcaacacacctcttaatgagagaatgatatctatg 2400  
|||||  
Db 5295 atgtgtaacatacagacgatatccggggcaacacacctcttaatgagagaatgatatctatg 5354  
QY 2401 gcccaagatgactgcgcggaaaaaacgcgatatgtgagcgacagcgccgcacttttagagcg 2460  
|||||  
Db 5355 gcccaagatgactgcgcggaaaaaacgcgatatgtgagcgacagcgccgcacttttagagcg 5414  
QY 2461 aagcctataagatattacaacacacgtgaagcgagatttgaggaaataaggaacgctg 2517  
|||||  
Db 5415 aagcctataagatattacaacacacgtgaagcgagatttgaggaaataaggaacgctg 5471  
|||||

RESULT 3  
AAS18307  
ID AAS18307 standard; DNA: 8509 BP.  
XX  
AC AAS18307;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DN DNA sequence of Ketogulonigenium endogenous plasmid pADM291.  
XX  
DE Cloning vector: Ketogulonigenium replicon; endogenous plasmid;  
XX  
KW transformed host cell; Escherichia coli; pADM291; circular; cyclic; ds.  
XX  
OS Ketogulonigenium sp. strain ADM291-19.  
XX  
PN WO200177347-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 05-APR-2001; 2001WO-US11059.  
XX  
PR 05-APR-2000; 2000US-194625P.  
XX  
PA (ARCH ) ARCHER-DANIELS MIDLAND CO.  
PA (DELI/) D'ELIA J.  
PA



XX D'Elia J;  
XX  
XX WPI: 2002-049150/06.  
XX  
XX Novel nucleic acid vector comprising Ketogulonigenium replicon found on  
XX a specific deposited endogenous plasmid, useful for producing  
XX polypeptides and/or transcripts by culturing host cells transformed  
XX with vector -  
XX  
XX Example 7; Fig 2; 66pp; English.  
XX  
XX The present invention relates to the isolation of vectors comprising  
XX a Ketogulonigenium replicon found on the endogenous plasmid, PADM291.  
XX The invention also describes methods of transforming host cells with  
XX the vectors and producing polypeptides and/or antisense transcripts by  
XX culturing the transformed host cells. The vectors are useful for  
XX transforming a host cell by conjugation or electroporation.  
XX The vectors which have a replicon functional in both Ketogulonigenium  
XX and Escherichia coli, enable the cloning of certain genes of  
XX Ketogulonigenium in E.coli as the latter is an efficient host for  
XX amplification of vector DNA. The present DNA sequence represents the  
XX Ketogulonigenium endogenous plasmid PADM291.  
XX  
XX Sequence 8509 BP; 2065 A; 2239 C; 2345 G; 1860 T; 0 other;

Query Match 100.0%; Score 2517; DB 24; Length 8509;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ggcattggtgcgaatactagatattgtgtgagtgctgtagcggtctgcagaaggtgt 60  
Db 2955 ggcattggtgcgaatactagatattgtgtgagtgctgtagcggtctgcagaaggtgt 3014  
Oy 61 ctgcgcgagagatctgtgtctcaggtlaaggcgacaatgagaggtgttagtgcacctg 120  
Db 3015 ctgcgcgagagatctgtgtctcaggtlaaggcgacaatgagaggtgttagtgcacctg 3074  
Oy 121 tatgctctctgctggtgcgcatgtgtcatctctgcgcgacatagatattccgctagag 180  
Db 3075 tatgctctctgctggtgcgcatgtgtcatctctgcgcgacatagatattccgctagag 3134  
Oy 181 gattactgatagttctcgtctcgtctgcggtctgcggtctgttcggtctgtgc 240  
Db 3135 gattactgatagttctcgtctcgtctgcggtctgcggtctgttcggtctgtgc 3194  
Oy 241 gggcctgtccctctgtccgcgtgtcctcaacttttcaacaatcaaaaatgggcgaagc 300  
Db 3195 gggcctgtccctctgtccgcgtgtcctcaacttttcaacaatcaaaaatgggcgaagc 3254  
Oy 301 ccttttgttctatagttcttatagttctatacgaataatacaataatataatagctt 360  
Db 3255 ccttttgttctatagttcttatagttctatacgaataatacaataatataatagctt 3314  
Oy 361 attcgcttaaaaggagataattggcgccgaagaaggagtaattggcgccgaagaaggat 420  
Db 3315 attcgcttaaaaggagataattggcgccgaagaaggagtaattggcgccgaagaaggat 3374  
Oy 421 aattggcgccgaagaaggagtaattggcgccgaatattcgttgtttaaattggggagatcc 480  
Db 3375 aattggcgccgaagaaggagtaattggcgccgaatattcgttgtttaaattggggagatcc 3434  
Oy 481 ccttaatacttctcccatgggaaagaacaacaagaatggcgccgacagccgtcttcgac 540  
Db 3435 ccttaatacttctcccatgggaaagaacaacaagaatggcgccgacagccgtcttcgac 3494  
Oy 541 cagacaaaactgtctccctcgcgaggtggcgagaggggtctatattgccaatccgcc 600  
Db 3495 cagacaaaactgtctccctcgcgaggtggcgagaggggtctatattgccaatccgcc 3554  
Oy 601 cgcctgcagcgcccaagctcatgtaattagatagcaactgcgggcgccgcatgtgct 660  
Db 601 cgcctgcagcgcccaagctcatgtaattagatagcaactgcgggcgccgcatgtgct 660

Db 3555 cgcctgcagcgcccaagctcatgtaattagatagcaactgcgggcgccgcatgtgct 3614  
Oy 661 gatgatgtgcgcatgaataatgctgctgcgcgacattcgcgcgaatcgaagcatgaataac 720  
Db 3615 gatgatgtgcgcatgaataatgctgctgcgcgacattcgcgcgaatcgaagcatgaataac 3674  
Oy 721 catgacgcgagagacgtctgcaccccgctgttcgagaagctagccactcgtgttgaccat 780  
Db 3675 catgacgcgagagacgtctgcaccccgctgttcgagaagctagccactcgtgttgaccat 3734  
Oy 781 gatgacctgcgaagaatgactcgtgacagtcgcgctgtgttcgaatgaaggcgcaatagac 840  
Db 3735 gatgacctgcgaagaatgactcgtgacagtcgcgctgtgttcgaatgaaggcgcaatagac 3794  
Oy 841 taccgcgaaggagcaagcgcggaactccttagtagagcttggaacttcggggtatattccgt 900  
Db 3795 taccgcgaaggagcaagcgcggaactccttagtagagcttggaacttcggggtatattccgt 3854  
Oy 901 cgtatggcgagagtcggaaccacttggtccatctcgaacgtcaaacggtatcacttc 960  
Db 3855 cgtatggcgagagtcggaaccacttggtccatctcgaacgtcaaacggtatcacttc 3914  
Oy 961 ggtagtaagatattcgtgtcgtgtgttcacagacgtctctagttccgcaattgtatcgg 1020  
Db 3915 ggtagtaagatattcgtgtcgtgtgttcacagacgtctctagttccgcaattgtatcgg 3974  
Oy 1021 atgagcgcaaaaaccttaccgttcccgagttgcggcgctcctctgagtgcccgagga 1080  
Db 3975 atgagcgcaaaaaccttaccgttcccgagttgcggcgctcctctgagtgcccgagga 4034  
Oy 1081 aagatgttcgtttggaacacgttcaacagattgtgctcacaacctgcaactggaatgagac 1140  
Db 4035 aagatgttcgtttggaacacgttcaacagattgtgctcacaacctgcaactggaatgagac 4094  
Oy 1141 aaccattatcgcgtctgacatgtacgcaagaacgcaacaagaatgtgcgctagcgtgcga 1200  
Db 4095 aaccattatcgcgtctgacatgtacgcaagaacgcaacaagaatgtgcgctagcgtgcga 4154  
Oy 1201 agtgtgactataggtcgtgggaatgaaagacgaccacaacgctgcgcggcgcgagctggcg 1260  
Db 4155 agtgtgactataggtcgtgggaatgaaagacgaccacaacgctgcgcggcgcgagctggcg 4214  
Oy 1261 ggttccaaggtcgtctcgaatgtctgtcgcgaaggggcgacgcaaaaacgataagacctcc 1320  
Db 4215 ggttccaaggtcgtctcgaatgtctgtcgcgaaggggcgacgcaaaaacgataagacctcc 4274  
Oy 1321 ttccagaagcgggcggaatcacctacgaatcgaatgtgtgtgtgagctggaacgtctgtct 1380  
Db 4275 ttccagaagcgggcggaatcacctacgaatcgaatgtgtgtgtgagctggaacgtctgtct 4334  
Oy 1381 ggcagcaaaagaagaacgatactgatacgtctcgaacttcggcggtttctgtcgggagaga 1440  
Db 4335 ggcagcaaaagaagaacgatactgatacgtctcgaacttcggcggtttctgtcgggagaga 4394  
Oy 1441 ggcgtgcgtctggaacgtctcaaacacgcaaaaactttttaaattctcgcgaagaagta 1500  
Db 4395 ggcgtgcgtctggaacgtctcaaacacgcaaaaactttttaaattctcgcgaagaagta 4454  
Oy 1501 ggggaaggtttgagttttgaggtatttcaacgcaatagtggtttaaagacttcgtgaaag 1560  
Db 4455 ggggaaggtttgagttttgaggtatttcaacgcaatagtggtttaaagacttcgtgaaag 4514  
Oy 1561 atgtgcaataatagcgttaagactatgaataacacggtcgaacggtctgcaaaaagcaacg 1620  
Db 4515 atgtgcaataatagcgttaagactatgaataacacggtcgaacggtctgcaaaaagcaacg 4574  
Oy 1621 gtttgcgcacgcgcaaacatcactcggcgctlaaaaagcggtaaaattccggtlaaaaag 1680  
Db 4575 gtttgcgcacgcgcaaacatcactcggcgctlaaaaagcggtaaaattccggtlaaaaag 4634  
Oy 1681 atgaactcggggacatgggttataatagatcccgagaattgcaacagatgttccctccatt 1740  
Db 4635 atgaactcggggacatgggttataatagatcccgagaattgcaacagatgttccctccatt 4694



QY 1741 caaagaataacaccgaacacctaacaacgagatataatgtaagcgtgataaacaacatg 1800  
|||  
DB 4695 caaagaataacaccgaacacctaacaacgagatataatgtaagcgtgataaacaacatg 4754  
QY 1801 aatgaactcagaatacgaacttagagcgtgagagtcgagacttcagagacttcatt 1860  
|||  
DB 4755 aatgaactcagaatacgaacttagagcgtgagagtcgagacttcagagacttcatt 4814  
QY 1861 ctgattccagagagagatcgcgcacaatgagcgcacatgagcgcgttcattcaattcat 1920  
|||  
DB 4815 ctgattccagagagagatcgcgcacaatgagcgcacatgagcgcgttcattcaattcat 4874  
QY 1921 caccgattgag 1980  
|||  
DB 4875 caccgattgag 4934  
QY 1981 cctggcttcag 2040  
|||  
DB 4935 cctggcttcag 4994  
QY 2041 aagtaataatagccacaataatagagagagagagagagagagagagagagagagagag 2100  
|||  
DB 4995 aagtaataatagccacaataatagagagagagagagagagagagagagagagagagag 5054  
QY 2101 caatacaagatattgatatcatcaatgagagagagagagagagagagagagagagagagag 2160  
|||  
DB 5055 caatacaagatattgatatcatcaatgagagagagagagagagagagagagagagagagag 5114  
QY 2161 cagcattgcttttctgattcttcaagtgcacaatacaatacgcgtgcgtccctaagaagatt 2220  
|||  
DB 5115 cagcattgcttttctgattcttcaagtgcacaatacgcgtgcgtccctaagaagatt 5174  
QY 2221 gtaactatgagagatctccacacttggtttcagacccgaataatgaagcgttgcagaagc 2280  
|||  
DB 5175 gtaactatgagagatctccacacttggtttcagacccgaataatgaagcgttgcagaagc 5234  
QY 2281 gttcttttgcag 2340  
|||  
DB 5235 gttcttttgcag 5294  
QY 2341 atggttaacatacacagacatcatccgggagcaacacaccctctatggaagatgagattatg 2400  
|||  
DB 5295 atggttaacatacacagacatcatccgggagcaacacaccctctatggaagatgagattatg 5354  
QY 2401 gcccgagagatctgag 2460  
|||  
DB 5355 gcccgagagatctgag 5414  
QY 2461 aagcctataacgagatcccaaaacactagaagccagatgagagagagagagagagagagag 2517  
|||  
DB 5415 aagcctataacgagatcccaaaacactagaagccagatgagagagagagagagagagagag 5471

## RESULT 4

AAS18306 standard; DNA; 2112 BP.

AAS18306;

12-MAR-2002 (first entry)

DNA sequence of Ketogulonigenium replicon from plasmid pADM291.

Cloning vector; Ketogulonigenium replicon; endogenous plasmid;

transformed host cell; Escherichia coli; pADM291; ds.

Ketogulonigenium sp. strain ADM291-19.

MO200177347-A2.

18-OCT-2001.

XX

XX

XX

XX

XX

XX

XX

XX

XX

PF 05-APR-2001; 2001WO-US11059.  
XX  
XX 05-APR-2000; 2000US-194625P.  
XX  
PA (ARCH) ARCHER-DANTELS MIDLAND CO.  
XX  
XX (DELLI) D'ELIA J.  
PI D'Elia J;  
XX  
XX WPI; 2002-049150/06.  
DR  
XX  
XX Novel nucleic acid vector comprising Ketogulonigenium replicon found on  
PT a specific deposited endogenous plasmid, useful for producing  
PT polypeptides and/or transcripts by culturing host cells transformed  
PT with vector  
XX  
PS Claim 2; Fig 1; 6pp; English.  
CC The present invention relates to the isolation of vectors comprising  
CC Ketogulonigenium replicon found on the endogenous plasmid, pADM291.  
CC The invention also describes methods of transforming host cells with  
CC the vectors and producing polypeptides and/or antisense transcripts by  
CC culturing the transformed host cells. The vectors are useful for  
CC transforming a host cell by conjugation or electroporation.  
CC The vectors which have a replicon functional in both Ketogulonigenium  
CC and Escherichia coli, enable the cloning of certain genes of  
CC Ketogulonigenium in E.coli as the latter is an efficient host for  
CC amplification of vector DNA. The present DNA sequence represents the  
CC replicon of Ketogulonigenium endogenous plasmid pADM291.  
XX  
SQ Sequence 2112 BP; 530 A; 509 C; 605 G; 468 T; 0 other;

Query Match 79.7%; Score 2007; DB 24; Length 2112;  
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 5; Indels 0; Gaps 0;  
Matches 2010; Conservative 0;

QY 1 ggcacatgggtcgaatcatagaatttctgtgaggtgcgtgacggtcgcagagagtg 60  
|||  
DB 1 ggcacatgggtcgaatcatagaatttctgtgaggtgcgtgacggtcgcagagagtg 60  
QY 61 ctggcgagagatccctcgtcctcagatgtagggcgacatgtagaggtgttagttgcccctg 120  
|||  
DB 61 ctggcgagagatccctcgtcctcagatgtagggcgacatgtagaggtgttagttgcccctg 120  
QY 121 tatcgtctctgctgtgagcattggtcattcctgcccgaacatagatatccgtctagag 180  
|||  
DB 121 tatcgtctctgctgtgagcattggtcattcctgcccgaacatagatatccgtctagag 180  
QY 181 gattactgatagtttctgctcgtcgtcgtgtgtcgtgtgtgtgtgtgtgtgtgtgtgtgtc 240  
|||  
DB 181 gattactgatagtttctgctcgtcgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtc 240  
QY 241 gggcgtgctccctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtc 300  
|||  
DB 241 gggcgtgctccctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtc 300  
QY 301 cctctgttctatagttcttatagttcttatagttcttatagttcttatagttcttatagttcttat 360  
|||  
DB 301 cctctgttctatagttcttatagttcttatagttcttatagttcttatagttcttatagttcttat 360  
QY 361 attcgtcttaaaaggagtaattggtgcgcaaaaggagtaattggtgcgcaaaaggagta 420  
|||  
DB 361 attcgtcttaaaaggagtaattggtgcgcaaaaggagtaattggtgcgcaaaaggagta 420  
QY 421 aattggccgcaaaaggagtaattggtgcgcaaaaggagtaattggtgcgcaaaaggagta 480  
|||  
DB 421 aattggccgcaaaaggagtaattggtgcgcaaaaggagtaattggtgcgcaaaaggagta 480  
QY 481 ccttaataatttcccatgag 540  
|||  
DB 481 ccttaataatttcccatgag 540



[illegible]

Db	1621	gtctgycgacccgcaacccatcctccgycgctaaaaacggtlaaaattccgyltaaaaag	1680
QY	1661	atgatacttgsggcaitggtgttatatgacccgagaaatltgacagagatgtttccctccatt	1740
Db	1661	atgaatcttgggcatctggtgttatatgacttcgcagaaatltgcacagagtgttccctccatt	1740
QY	1741	caagaatacacccgaanaaccccttaacacgcaatatactgtgtagcgtatgtaaacacag	1800
Db	1741	caagaatacacacccgaacaccccttaacacgcaatatactgtgtagcgtatgtaaacacag	1800
QY	1801	aaatgacctcaagaatactagcgcatltgagcgttgaagttcggaccttaacgcatgtcttat	1860
Db	1801	aaatgacctcaagaatactagcgcatltgagcgttgaagttcggaccttaacgcatgtcttat	1860
QY	1861	ctgactgcccagagagatcgcgcacaaatltgycgcacatgtgcccagcgccttccaatttcatt	1920
Db	1861	ctgactgcccagagagatcgcgcacaaatltgycgcacatgtgcccagcgccttccaatttcatt	1920
QY	1921	caaccgatgagagagagaaagacccgccccctccaanaaacaagaatltggtgagaatattctgat	1980
Db	1921	caaccgatgagagagagagagacccgccccctccaanaaacaagaatltggtgagaatattctgat	1980
QY	1981	cctgagcttcaagagaccttgcctttaaacttgaa	2015
Db	1981	cctgagcttcaagagaccttgcctttaaacttgaa	2015

RESULT	5
AA518308	standard; DNA; 5859 BP.
ID	AA518308
XX	
AC	AA518308;
XX	
DT	12-MAR-2002 (first entry)
XX	
DE	DNA sequence of shuttle vector plasmid pADM291-4.
XX	
XX	Cloning vector; Ketogulonigenium replicon; endogenous plasmid; mutant;
RW	transformed host cell; Escherichia coli; pADM291-4; circular; cyclic;
XX	shuttle vector plasmid; ds.
XX	
OS	Ketogulonigenium sp. strain ADM291-19.
OS	Synthetic.
XX	
PN	WO200177347-A2.
XX	
PD	18-OCT-2001.
XX	
PF	05-APR-2001; 2001WO-US11059.
XX	
PR	05-APR-2000; 2000US-194625P.
XX	
XX	(ARCH ) ARCHER-DANIELS MIDLAND CO.
PA	(DELI/) D'ELIA J.
XX	
PI	D'Elia J;
XX	
DR	WPI; 2002-049150/06.
XX	
PT	Novel nucleic acid vector comprising Ketogulonigenium replicon found on
PT	a specific deposited endogenous plasmid, useful for producing
PT	polypeptides and/or transcripts by culturing host cells transformed
PT	with vector -
XX	
PS	Claim 3; Fig 3; 66pp; English.
XX	
XX	The present invention relates to the isolation of vectors comprising
CC	a Ketogulonigenium replicon found on the endogenous plasmid, pADM291.
CC	The invention also describes methods of transforming host cells with
CC	the vectors and producing polypeptides and/or antisense transcripts by
CC	culturing the transformed host cells. The vectors are useful for
CC	transforming a host cell by conjugation or electroporation.



CC The vectors which have a replicon functional in both *Ketogulonigenium*  
CC and *Escherichia coli*, enable the cloning of certain genes of  
CC *Ketogulonigenium* in *E. coli* as the latter is an efficient host for  
CC amplification of vector DNA. The present DNA sequence represents the  
CC shuttle vector plasmid pADM291-4.

XX Sequence 5859 BP; 1444 A; 1487 C; 1566 G; 1362 T; 0 other;

Query Match 79.7%; Score 2007; DB 24; Length 5859;  
Best Local Similarity 99.88; Pred. No. 0;  
Matches 2010; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggcacatggtcgaataatcatagaaatttctgtgaggtgctgacgctcgcagacaggtc 60  
DB 402 ggcacatggtcgaataatcatagaaatttctgtgaggtgctgacgctcgcagacaggtc 461  
QY 61 ctgcgcgagatctctgtctcgaagtagggcgacatgagaggtgttaattgccccctg 120  
DB 462 ctgcgcgagatctctgtctcgaagtagggcgacatgagaggtgtgttaattgccccctg 521  
QY 121 tatcgctctcgtgcgcgcatgtgcatccgtccgcgacatatgtatctcgtagag 180  
DB 522 tatcgctctcgtgcgcgcatgtgcatccgtccgcgacatatgtatctcgtagag 581  
QY 181 gattactgaattcttcgctcgtgcgtgtcgcgctgtcgcgctgtcgcgctgtgc 240  
DB 582 gattactgaattcttcgctcgtgcgtgtcgcgctgtcgcgctgtcgcgctgtgc 641  
QY 241 gggcgtgtccctctgtccgcgctgtcctcaattttcaacaatacaaaaatggcgagagc 300  
DB 642 gggcgtgtccctctgtccgcgctgtcctcaattttcaacaatacaaaaatggcgagagc 701  
QY 301 cctctctgtctatagtcttctatagtctatgaatgaataatcatatcatatagctt 360  
DB 702 cctctctgtctatagtcttctatagtctatgaatgaataatcatatcatatagctt 761  
QY 361 attcgcttaaaaggagtaatttggccgcaaaaggagtaatttggccgcaaaaggaggt 420  
DB 762 attcgcttaaaaggagtaatttggccgcaaaaggagtaatttggccgcaaaaggaggt 821  
QY 421 aattgggcgcgaagagagtaatttggccgcaaaaggagtaatttggccgcaaaaggaggt 480  
DB 822 aattgggcgcgaagagagtaatttggccgcaaaaggagtaatttggccgcaaaaggaggt 881  
QY 481 ccttaataattctcccatgtggaagacaacacaaagtggccgcgacccgcttcgac 540  
DB 882 ccttaataattctcccatgtggaagacaacacaaagtggccgcgacccgcttcgac 941  
QY 541 cagacaaaaactgtgtccctcgcgcgaggttggcgaaggggtctatagcgaatccgccc 600  
DB 942 cagacaaaaactgtgtccctcgcgcgaggttggcgaaggggtctatagcgaatccgccc 1001  
QY 601 cgcctcgaagcgtcaagctcatgcatthaatgataagccaactgcggcgcgcgataggtc 660  
DB 1002 cgcctcgaagcgtcaagctcatgcatthaatgataagccaactgcggcgcgcgataggtc 1061  
QY 661 gatgatgtgcgcataaattgagctggccgacatctgcgcgaatcgagatgaaaac 720  
DB 1062 gatgatgtgcgcataaattgagctggccgacatctgcgcgaatcgagatgaaaac 1121  
QY 721 catgacgctgagagcctgcgacccgctgtcgaagagctagccgcgtgctgtgacat 780  
DB 1122 catgacgctgagagcctgcgacccgctgtcgaagagctagccgcgtgctgtgacat 1181  
QY 781 gatgacccctgcaaaatgatcgtgacagatgcgcgctgtgcatgagcgcgaaatagac 840  
DB 1182 gatgacccctgcaaaatgatcgtgacagatgcgcgctgtgcatgagcgcgaaatagac 1241  
QY 841 tacgcgcagagagcgaagcgcgaaactctcagtgagcgttgagccttcgcggagctacatccgt 900  
DB 1242 tacgcgcagagagcgaagcgcgaaactctcagtgagcgttgagccttcgcggagctacatccgt 1301

QY 901 cgtatgctgcgcgagtagtcgaacacactggtgcatcttcctgacccgtcaaacgtaattccatctc 960  
DB 1302 cgtatgctgcgcgagtagtcgaacacactggtgcatcttcctgacccgtcaaacgtaattccatctc 1361  
QY 961 ggtatgaatgattccgtctcgtctgtcttcacagcagctctctagtcttcgcaatcttgatcgc 1020  
DB 1362 ggtatgaatgattccgtctcgtctgtcttcacagcagctctctagtcttcgcaatcttgatcgc 1421  
QY 1021 atgagcgcgaaaaccttacggtcccccaggttgcggcgctccctctgaggttcccgagagga 1080  
DB 1422 atgagcgcgaaaaccttacggtcccccaggttgcggcgctccctctgaggttcccgagagga 1481  
QY 1081 aagatggttgcgttggaacgcagcttaacagattgtctctcaaacctgcgaactggaatgagtc 1140  
DB 1482 aagatggttgcgttggaacgcagcttaacagattgtctctcaaacctgcgaactggaatgagtc 1541  
QY 1141 aaccttatacgtctgacattgagcgcgaagccgaacccaagattggccgtagctgtgca 1200  
DB 1542 aaccttatacgtctgacattgagcgcgaagccgaacccaagattggccgtagctgtgca 1601  
QY 1201 agtctgactatagagcttggaagtagtaagacgacccaacgctgcgcagcgcgagctggcg 1260  
DB 1602 agtctgactatagagcttggaagtagtaagacgacccaacgctgcgcagcgcgagctggcg 1661  
QY 1261 ggtccaaagctcgtcgaagatgctctgcgcagagagggcgagcgaacagatagccctcc 1320  
DB 1662 ggtccaaagctcgtcgaagatgctctgcgcagagagggcgagcgaacagatagccctcc 1721  
QY 1321 ttccccaagagcgggcgagatgctctgcagagaggttgcgtgagagctgtgaaacgctctgct 1380  
DB 1722 ttccccaagagcgggcgagatgctctgcagagaggttgcgtgagagctgtgaaacgctctgct 1781  
QY 1381 ggaagacaagaagacaacgactgtatcgctcctcaagctccggcgttctctgcgcggaagga 1440  
DB 1782 ggaagacaagaagacaacgactgtatcgctcctcaagctccggcgttctctgcgcggaagga 1841  
QY 1441 ggcgtgtgctctgacgctgcaaacatcgcgaacaaactgttttagattcttcgcgcaaaagta 1500  
DB 1842 ggcgtgtgctctgacgctgcaaacatcgcgaacaaactgttttagattcttcgcgcaaaagta 1901  
QY 1501 gggagagttgagtttgaggtatttcacccgcgaatggtgttaatgactcttcgtgaaacg 1560  
DB 1902 gggagagttgagtttgaggtatttcacccgcgaatggtgttaatgactcttcgtgaaacg 1961  
QY 1561 atgtgcataatagcgtgaagactatgaataacacagcctgcgagcgtcgaagaagcaacgg 1620  
DB 1962 atgtgcataatagcgtgaagactatgaataacacagcctgcgagcgtcgaagaagcaacgg 2021  
QY 1621 gttgtgcgacccgcgaacatcaactcgtggtcgttaaaacggttaaatcttcgcggttaaaaag 1680  
DB 2022 gttgtgcgacccgcgaacatcaactcgtggtcgttaaaacggttaaatcttcgcggttaaaaag 2081  
QY 1681 atgaaatctggggcatggtgttaataatcttcgagaattgcagagatgtttccctccattt 1740  
DB 2082 atgaaatctggggcatggtgttaataatcttcgagaattgcagagatgtttccctccattt 2141  
QY 1741 caaagaataacacgcgaacacactaacacgcgaacgaagtaatgtgaagcgtgataaacacatg 1800  
DB 2142 caaagaataacacgcgaacacactaacacgcgaacgaagtaatgtgaagcgtgataaacacatg 2201  
QY 1801 aaatgaccccaagaatacagcgcatlagagcgttgaagttcggactttagcgatgctttat 1860  
DB 2202 aaatgaccccaagaatacagcgcatlagagcgttgaagttcggactttagcgatgctttat 2261  
QY 1861 ctgatccgaaggagatcgcgcgaacaaatgvcgcgaatgvcgcgaagcgtcttcaattcat 1920  
DB 2262 ctgatccgaaggagatcgcgcgaacaaatgvcgcgaatgvcgcgaagcgtcttcaattcat 2321  
QY 1921 caccgataagagagagagacgcgcgcctcacaacaaagaatgtgtgagaatatcttcat 1980  
DB 2322 caccgataagagagagagacgcgcgcctcacaacaaagaatgtgtgagaatatcttcat 2381  
QY 1981 ccttggtctcaggaagccttgctttaaactctgaa 2015











Db 8696 AAA 8694

RESULT 10

ABL34011

AC ABL34011;

XX 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1984.

KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianemic; cytosolic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

PE 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPICENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation

PS Claim 1; SEQ ID NO 1984; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.

XX Sequence 6050 BP; 2099 A; 34 C; 959 G; 2958 T; 0 other;

Query Match 1.5%; Score 39; DB 24; Length 6050;  
Best Local Similarity 61.2%; Pred. No. 1.7;  
Matches 63; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

OY 2094 atgtatcatcatcaatgattgatattcatcaatggagagaattacatgtatcacagg 2153

DB 2977 atgtaatgtttataaagtttgagatttgaatggagagaatttaaatatattataa 3036

OY 2154 atcatcacagcattgttttgatttctcaagtgcatacataa 2196

DB 3037 atgataattatatttttttaattttgtttttaagtaa 3079

RESULT 11

AAH24065/C

ID AAH24065 standard; DNA; 4590 BP.

XX AAH24065;

DT 29-AUG-2001 (first entry)

DE Yeast AOD9604-associated DNA sequence, SEQ ID NO:1.

KW Human growth hormone analogue peptide; hGH; AOD9604; lipid metabolism;  
KW modulation; lipolysis stimulation; hormone-sensitive lipase stimulation;  
KW lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;  
KW functional food; transgenic yeast; fat/lean ratio; food use; ds.

OS Saccharomyces cerevisiae.

PH Key Location/Qualifiers

FT misc\_feature 10

FT /tag= a "Represented as \* in the specification"

FT /note= 3617

FT /tag= b "Represented as \* in the specification"

FT /note= 3649

FT /tag= c "Represented as \* in the specification"

FT /note= 3679

FT /tag= d "Represented as \* in the specification"

FT /note= 3819

FT /tag= e "Represented as \* in the specification"

FT /note= 3862

FT /tag= f "Represented as \* in the specification"

FT /note= 3864

FT /tag= g "Represented as \* in the specification"

FT /note= 3888

FT /tag= h "Represented as \* in the specification"

FT /note= 3890

FT /tag= i "Represented as \* in the specification"

FT /note= 3912

FT /tag= j "Represented as \* in the specification"

FT /note= 3914

FT /tag= k "Represented as \* in the specification"

FT /note= 3938

FT /tag= l "Represented as \* in the specification"

FT /note= 3939

FT /tag= m "Represented as \* in the specification"

FT /note= 3941

FT /tag= o "Represented as \* in the specification"

FT /note= 3943

FT /tag= p "Represented as \* in the specification"

FT /note= 4361

FT /tag= q "Represented as \* in the specification"

XX WO200133977-A1.

XX 17-MAY-2001.

XX 06-NOV-2000; 2000WO-AU01362.

XX 05-NOV-1999; 99AU-0003875.

XX (META-) METABOLIC PHARM LTD.







RESULT 13  
 ID AA161373/c  
 AA161373 standard; DNA; 513445 BP.  
 AC  
 AA161373; ~  
 DT 16-OCT-2001 (first entry)  
 DE Soybean 318013 region A3, SEQ ID NO: 4.  
 KW Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;  
 KW SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;  
 KW 240017 region G3; 318013 region A3; 515002 region G2; ds.  
 OS Glycine max.  
 PN WO200151627-A2.  
 PD 19-JUL-2001.  
 PF 05-JAN-2001; 2001WO-US00552.  
 PR 07-JAN-2000; 2000US-0174880.  
 PA (MONS ) MONSANTO CO.  
 PI Haughe BM, Wang ML, Parsons JD, Parnell LD;  
 DR WPI: 2001-425872/45.  
 DR P-FSDB: AAM42216.  
 PS  
 PT New purified nucleic acid for producing a soybean plant having soybean  
 PT cyst nematode resistance and for use in plant breeding programs -  
 PS Claim 30; Page 596-893; 1353pp; English.  
 CC The invention relates to nucleic acid molecules from regions of the  
 CC soybean genome which are associated with soybean cyst nematode (SCN)  
 CC resistance. The nucleic acids are used to transform plants, and can  
 CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.  
 CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes  
 CC of soybean plants and for introgressing SCN resistance or partial SCN  
 CC resistance into soybean plants. They can also be used in plant breeding  
 CC programmes. The invention also relates to proteins encoded by such  
 CC nucleic acid molecules, as well as antibodies capable of recognising  
 CC these proteins. The present sequence is a nucleic acid molecule  
 CC provided in the specification.  
 QQ Sequence 513445 BP; 173367 A; 85402 C; 83912 G; 170492 T; 272 other;

Query Match	1.5%	Score 37.8	DB 22	Length 513445
Best Local-Similarity	48.4%	Pred. No. 61		
Matches 105	Conservative 0	Mismatches 112	Indels 0	Gaps 0
QY 2004	ttaaaaccttgatcagcattctctagcatgctgataagaataataatagccacaatagag	2063		
Db 410832	TAACAAATATGCAATGAGAGGAATTTATTCATATATATATATTAATTTGCTAAATAATCAAG	410773		
QY 2064	cggccatttctcattcacacataacagctctcctcaatgatcaataatcaagatgatattct	2123		
Db 410772	TTAACTATAGTTTATATATATAGTATATAGTTGATGATGTTGTTTTTATTTATATTTTTC	410713		
QY 2124	caatgagagaagaatttacaatgataacaggaatcacacacagaatttgtttctatttcta	2183		
Db 410712	ATATGATATACATATAGAAATTAATCAATGATTCGATTCGATTCATCTGTTATTTTATTTTCTA	410653		
QY 2184	agtcctaacataactatcgcgcgccttaagaagaatt	2220		
Db 410652	AATTCAAATGTGCCCTATATATATATCAACCCAAAAAATGTT	410616		

## RESULT 14

AAAF26561/C  
ID AAAF26561 standard; DNA; 4043 BP.  
XX  
AC AAAF26561;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE DNA encoding human secreted protein #15.  
XX  
KW Secreted protein; gene therapy; Vaccine; cancer; leukemia;  
KW autoimmune disease; allergy; inflammation; graft rejection;  
KW hyperproliferation; cardiovascular; infection; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200076531-A1.  
XX  
PD 21-DEC-2000.  
XX  
PF 01-JUN-2000; 2000WO-US15137.  
XX  
PR 11-JUN-1999; 99US-0138625.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Komatsoulis GA;  
DR WPI; 2001-071148/08.  
XX  
PT Nucleic acids encoding 47 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
PT disease and diabetic retinopathy -  
XX  
PS Claim 1; Page 449-454; 525pp; English.  
XX  
CC The present invention relates to 26 secreted human proteins. The  
CC proteins may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate polypeptide expression.  
CC For example, they may be used in gene therapy or in vaccines.  
CC Typical of diseases which are potentially treatable are cancers  
CC (including leukemia), autoimmune diseases, allergies, inflammation,  
CC graft rejection, hyperproliferation, cardiovascular diseases  
CC (particularly critical limb ischemia and coronary disease) and any  
CC involving abnormal angiogenesis, neurodegeneration and/or  
CC infectious diseases.  
XX  
Sequence 4043 BP; 1329 A; 695 C; 694 G; 1324 T; 1 other;

	Query Match	Best Local Similarity	1.5% 50.0% Pred. No. 3.4;	Score 37.6;	DB 22;	Length 4043;
	Matches	94	Conservative	0;	Mismatches	94;
					Indels	0;
					Gaps	0.
QY	2003	ttlaaacctgaatcagcatctctagcgaibctgataagaagtaaalatagccacatataga				2062
Db	2832	TTTTTATTAAATATCTAGAGATCTAAGATCTTCTTAAATACCAATTTACCAACAAATATA				2763
QY	2063	gcggccatcttcacatcagcatcactatcctatgcatcaatataaagattgatcatca				2122
Db	2762	TCATCAAAAATATTTACTCAATTCCTGCGCCAGGATTCGTGGATGACCAACACTTAACTTCT				2703
QY	2123	tcaatggaagaagaattcatatgatacacagaagatcaacagatattgttttgatattct				2182
Db	2702	TACTTAGCAATGATATATATATTTATAAGATATCCATATAAAGATCTTGTCTTTAAT				2643
QY	2183	aagtcgcta	2190			
Db	2642	GAAAACCTA	2635			

RESULT	15
AAF64182/c	
ID	AAF64182 standard; cDNA; 4044 BP.



Sequence 4044 BP: 1329 A; 695 G; 695 G; 1323 T; 2 other;

Query Match	1.58;	Score 37.6;	DB 22;	Length 4044;
Best Local Similarity	50.08;	Pred. No. 3.4;		
Matches 94;	Conservative	0;	Mismatches 94;	Indels 0;
			Gaps	0;

QY	2003	ttaaaccttgatcagcattcttagcgatgcgtgataagaaaglaataatagcccaatga	2062
Db	2832	TTTTATTATATATATGAGATCTGAGATCTCTTATATTAACCTATATTCACACAAATATAA	2763
QY	2063	ggggcatttccatccatcaacagctcgcattcgtatgcattcaatcaagaattgatattca	2122
Db	2762	TGATCAAAATATATTACTCAATGTTGTCGCCAGAGATGTGTCGATPACCAACATTAATCTTC	2703
QY	2123	tcaatgagagaatattcttcattgattcgcagatgcacagcagcattggttttgattctt	2182
Db	2702	TACTTGCAATGATATATATTATTATTAAGATATGCCATATATAAAGATCTTGTCTTTATAT	2643
QY	2183	aagtgcata	2190
Db	2642	GAAAACTA	2635

Search completed: August 20, 2002, 21:33:26  
Job time: 5893 sec



**THIS PAGE BLANK (USPTO)**

---



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 19:48:08 ; Search time 62.61 Seconds  
(without alignments)  
9874.773 Million cell updates/sec

Title: US-09-826-206-4  
Perfect score: 2517  
Sequence: 1 ggcgaatgggtcgaattcat.....tgaggaataatagggaactgt 2517

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_patents\_NA:\*
- 1: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/1na/PTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/2/1na/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	45.8	1.8	7218	1 US-08-232-463-14	Sequence 14, Appl
2	37.4	1.5	7218	1 US-08-232-463-14	Sequence 14, Appl
3	36.6	1.5	6216	3 US-09-213-053-1	Sequence 1, Appl
4	36	1.4	289	4 US-09-007-005-17	Sequence 17, Appl
5	36	1.4	289	4 US-09-244-796-17	Sequence 17, Appl
C 6	35.8	1.4	4403765	4 US-09-103-840A-2	Sequence 2, Appl
C 7	35.8	1.4	4411529	4 US-09-103-840A-1	Sequence 1, Appl
C 8	35.4	1.4	857	5 PCT-US95-17083-9	Sequence 9, Appl
C 9	35.4	1.4	920	5 PCT-US95-17083-7	Sequence 7, Appl
C 10	35.4	1.4	975	5 PCT-US95-17083-5	Sequence 5, Appl
C 11	35.4	1.4	1104	5 PCT-US95-17083-3	Sequence 3, Appl
C 12	35.4	1.4	1167	5 PCT-US95-17083-1	Sequence 1, Appl
C 13	35.4	1.4	2471	1 US-08-444-231-18	Sequence 18, Appl
C 14	35.4	1.4	2471	1 US-08-152-443A-18	Sequence 18, Appl
C 15	35.4	1.4	2534	2 US-08-219-237B-1	Sequence 1, Appl
C 16	35.4	1.4	2534	4 US-08-468-560C-1	Sequence 16, Appl
C 17	35.4	1.4	2534	4 US-09-180-100-16	Sequence 16, Appl
C 18	35.4	1.4	2551	4 US-09-280-640-1	Sequence 1, Appl
C 19	35.2	1.4	87350	3 US-08-781-891-79	Sequence 79, Appl
C 20	34.4	1.4	500	3 US-09-141-000-2	Sequence 2, Appl
C 21	34.4	1.4	5176	2 US-09-182-024A-1	Sequence 1, Appl
C 22	34.2	1.4	1643	2 US-08-486-839-3	Sequence 3, Appl
C 23	34.2	1.4	1643	3 US-09-151-011-3	Sequence 3, Appl
C 24	34.2	1.4	1643	4 US-09-343-623-3	Sequence 3, Appl
C 25	34.2	1.4	1713	2 US-08-486-839-5	Sequence 5, Appl
C 26	34.2	1.4	1713	3 US-09-151-011-5	Sequence 5, Appl
C 27	34.2	1.4	1713	4 US-09-343-623-5	Sequence 5, Appl

C 28	33.6	1.3	2756	1 US-08-187-793-1	Sequence 1, Appl
C 29	33.6	1.3	3318	1 US-08-187-793-3	Sequence 3, Appl
C 30	33.6	1.3	4104	1 US-07-998-003A-94	Sequence 94, Appl
31	33.6	1.3	4104	1 US-08-453-224B-94	Sequence 94, Appl
32	33.6	1.3	4104	1 US-08-453-695A-94	Sequence 94, Appl
33	33.6	1.3	4104	1 US-08-268-161A-94	Sequence 94, Appl
34	33.6	1.3	4104	2 US-08-453-702A-94	Sequence 94, Appl
35	33.6	1.3	4104	1 US-09-099-639-94	Sequence 94, Appl
36	33.6	1.3	4104	5 PCT-US95-12588-94	Sequence 94, Appl
37	33.6	1.3	4104	5 PCT-US95-08071-94	Sequence 94, Appl
38	33.6	1.3	4650	1 US-07-998-003A-102	Sequence 102, App
39	33.6	1.3	4650	1 US-08-453-274B-102	Sequence 102, App
40	33.6	1.3	4650	1 US-08-453-695A-102	Sequence 102, App
41	33.6	1.3	4650	1 US-08-268-161A-102	Sequence 102, App
42	33.6	1.3	4650	2 US-08-453-702A-102	Sequence 102, App
43	33.6	1.3	4650	4 US-09-099-639-102	Sequence 102, App
44	33.6	1.3	4650	5 PCT-US93-12588-102	Sequence 102, App
45	33.6	1.3	4650	5 PCT-US95-08071-102	Sequence 102, App

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
: Sequence 14, Application US/08232463  
: Patent No. 5670367  
: GENERAL INFORMATION:  
: APPLICANT: DORNER, F.  
: APPLICANT: SCHEIFLINGER, F.  
: APPLICANT: FALKNER, F. G.  
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
: NUMBER OF SEQUENCES: 52  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Foley & Lardner  
: STREET: 1800 Diagonal Road, Suite 500  
: CITY: Alexandria  
: STATE: VA  
: COUNTRY: USA  
: ZIP: 22313-0299  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/232,463  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US/07/935,313  
: FILING DATE:  
: APPLICATION NUMBER: EP 91 114 300.6  
: FILING DATE: 26-AUG-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: BENT, Stephen A.  
: REGISTRATION NUMBER: 29,768  
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
: TELEPHONE: (703)836-9300  
: TELEFAX: (703)683-4109  
: TELEX: 899149  
: INFORMATION FOR SEQ ID NO: 14:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 7218 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: IMMEDIATE SOURCE:  
: CLONE: pTZgpt-F15  
: US-08-232-463-14



[illegible]

2  
 US-08-232-463-14  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHETTLINGER, F.  
 ; APPLICANT: FALKNER, F. G.  
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,463  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/935,313  
 ; FILING DATE:  
 ; APPLICATION NUMBER: EP 91 114 300.6  
 ; FILING DATE: 26-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703)836-9300  
 ; TELEFAX: (703)683-4109  
 ; TELEX: 899149  
 ; INFORMATION FOR SEQ ID NO: 14:

```

;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 7218 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;     CLONE: pTigpt-F15
;
US-08-232-463-14

```

Query Match	1.5%;	Score 37.4;	DB 1;	Length 7218;
Best Local Similarity	9.3%;	Pred. No.0.48;	117;	Indels 0;
Matches	26;	Conservative 136;	Mismatches 117;	Gaps 0;
QY	111	ttgcgccctgtctgcctctctgcgttgagcattgagtcacccgcagacatgatat	170	
Db	1213	yy	1272	
QY	171	tcgcctagagagataactagatctgtcctcgcgcgagcttgcggctctcgggcttgc	230	
Db	1273	yy	1332	
QY	231	cggagcttgcggcctgtccctcctgtccgcgcgttcctcactttccacataaaaaa	290	
Db	1333	yy	1392	
QY	291	tgggcgaaagccctctgtctctatagttcttaagtatcaagaaattcacataatta	350	
Db	1393	yygTACAAATCTTGTATC	1452	
QY	351	tcaatagcttattcgtcttaaaaggaagtaattgggcgcg	389	
Db	1453	TCTTTAACTACTGCATAGATAGTAAATTAAGTGATGC	1491	

```

RESULT 3
; Sequence 1, Application us/09213053
; Patent No. 6159477
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; APPLICANT: BAUDOU, Philippe
; TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
; TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
; TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
; FILE REFERENCE: 454313-2200
; CURRENT APPLICATION NUMBER: US/09/213,053
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 9608242
; EARLIER FILING DATE: 1996-06-27
; EARLIER APPLICATION NUMBER: PCT/FR97/01115
; EARLIER FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6216
; TYPE: DNA
; ORGANISM: Canine herpesvirus
; US-09-213-053-1

```

	Query Match	1.5%;	Score 36.6;	DB 3;	Length 6216;
	Best Local Similarity	47.2%;	Pred. NO. 0.78;	Mismatches 124;	Indels 0;
	Matches 11;	Conservative			Gaps 0.
QY	1994	agccctgaccttaaacccggaatcagcactctagcagtcgcatagaagaaatatagc	2053		
Db	3096	atcatttccataataatattgctctggatcgcgactctggtctaccatgtatgtatgt	3155		
QY	2054	cacaaatagagcgccatttccattcacatacagaatcatcatgtatgataataaagtat	2113		
Db	3156	aaccacaataatcactcttatgtatcttgacttcaagctcgaatttaacagccttaatat	3215		







Query Match	1.4%	Score 35.4	DB 5	Length 920
Best Local Similarity	50.9%	Pred. No. 0.59	81	Indels 0
Matches 84	Conservative	0	Mismatches	Gaps 0
QY 2109	agattgatattcaatcaatgagaagaattacatglatcacagatcatcaagcattt	2168		
Db 864	AGATTTTACGGCAGCGATTTAGAATCTTTTCAACACTAATTCATATCTCGAAGCTG	805		
QY 2169	gttttctatcttctaagtcctaacataactatcgctggccctaagaagaattgactatt	2228		
Db 804	AATTTGTGTTTTCCTCTAGCAGCAAGCTTTGGATTTTCATTTTCGAAAGTTGAATTTTC	745		
QY 2229	gcagatcctacaccttgggtttcagaccgataattcagctttgtc	2273		
Db 744	TGAATCACTAGTATGTCCTTGAGGATGATAGTCTGAATTTTCTC	700		

RESULT 10  
PCT-0595-17083-5/c



```

: Sequence 5 Application PC/TUS9517083
:
: GENERAL INFORMATION:
:
: APPLICANT:
:
: TITLE OF INVENTION: SECRETED HUMAN FAS ANTIGEN
:
: NUMBER OF SEQUENCES: 16
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: PCT/US95/17083
:
: FILING DATE: CONCURRENTLY HERewith
:
: CLASSIFICATION:
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: US95 08/371,263
:
: FILING DATE: 23-DEC-1994
:
: INFORMATION FOR SEQ ID NO: 5:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 975 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: PCT-US95-17083-5

```

Query Match	1.4%	Score 35.4	DB 5	Length 975
Best Local Similarity	50.9%	Pred. No. 0.61		
Matches	84	Conservative	0	Mismatches 81; Indels 0; Gaps 0;
QY 2109	agattgatactcatcgaatgagaagaattatcatgatacagagatcatcacagcatlt	2168		
Db 919	AGTATTTCACACGACGCTTTAAGAATCTTTCAACACTAATTGCATTACTACACACTG	860		
QY 2169	gtctttgattctctaagtgcatacaatacctgcctgcccctaagaagatgtgactat	2228		
Db 859	AATTGTGTTTTCACCTAGACCAACCTTTGGATTTTCATTTTCGAAGTTTGAAATTTTC	800		
QY 2229	gcagatctcaaccttgggtttcagacgcgataatcacgcttgc	2273		
Db 799	TGAGCACTAGTAATGTCCTTGAGGATATATGCTGAATTTTCTC	755		

```

RESULT 11
PCT-US95-17083-3/c
Sequence 3, Application PC/TUS9517083
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: SECRETED HUMAN FAS ANTIGEN
NUMBER OF SEQUENCES: 16
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17083
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/371,263
FILING DATE: 23-DEC-1994
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1104 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-17083-3

```

Query Match	1.4%	Score 35.4	DB 5	Length 1104
Best Local Similarity	50.9%	Pred. No. 0.66		
Matches	84	Conservative	0	Matches 81
				Indels 0
				Gaps 0
Qy 2109	agatgatcattc	atcatgtagaagaatt	catatcatgatacaagatcatcacaagcattt	2168
Db 1048	AGATATTACACG	CCACTTTTAAGAAVCTTTTCAACACTAACTAAATTTGGATATATACCAACAACG		989
Oy 2169	gtttttgattctc	aagtcgaacataactatcgcgtgcgcctcaagaagattgtactatt	2228	
Db 988	AATTTGTTGTTTTCAC	CTCTAGACCCACACTTTGGATTCATTTCATTTCTGAGATTTTGAATTTTC	929	

Oy 2229 gcaglatctcaccttggttcgaccgataattacagcttgc 2273  
||| || | ||| | ||| ||  
Db 928 TGAGTCACTAGTAATGTCCCTTGAGGATGATAGTCTGAATTTTC 884

```

1 RESULT 12
2 PCT-US95-17083-1/c
3 : Sequence 1, Application PC/TUS9517083
4 : GENERAL INFORMATION:
5 : APPLICANT:
6 : TITLE OF INVENTION: SECRETED HUMAN FAS ANTIGEN
7 : NUMBER OF SEQUENCES: 16
8 : CURRENT APPLICATION DATA:
9 : APPLICATION NUMBER DATA: PCT/US95/17083
10 : FILING DATE: CONCURRENTLY HEREWITH
11 : CLASSIFICATION:
12 : PRIOR APPLICATION DATA:
13 : APPLICATION NUMBER: USSN 08/371,263
14 : FILING DATE: 23-DEC-1994
15 : INFORMATION FOR SEQ ID NO: 1:
16 : SEQUENCE CHARACTERISTICS:
17 : LENGTH: 1167 base pairs
18 : TYPE: nucleic acid
19 : STRANDEDNESS: single
20 : TOPOLOGY: linear
21 PCT-US95-17083-1

```

Query Mch	1.48	Score 35.4	DB 5	Length
Best Local Similarity	50.98	Pred. No. 0.68		
Matches 84	Conservative	0	Mismatches 81	Indel
Qy 2109	agatcgaatcatcaagaagaattacatgcatcacaggaatc			
Db 1111	AGTATTACACCCAGCGATTAAAGAACTTTTCAAACTAATTCATATAT			
Qy 2169	gttttcgtatttctaagtcctaataactatcgctgcccataaaga			
Db 1051	AATTGTGTTTTCACATCTAGACCAAGCTTGGATTCTTCAGAGCT			
Qy 2229	gcagatctcaacctggctttgcagacgataatlaagccttgtc 2273			
Db 991	TGAGCACTACTAATGTCCTTGAGAGATAGTCGTAAATTTTTC 947			

RESULT 13  
; Sequence 18, Application US/08444231  
; Patent No. 5652210  
; GENERAL INFORMATION:  
; APPLICANT: BARR, PHILIP J.  
; APPLICANT: SHAPIRO, JOHN P.  
; APPLICANT: KIEFER, MICHAEL C.  
; TITLE OF INVENTION: NOVEL FAS PROTEIN AND METHODS OF U  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,231  
; FILING DATE: 18-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:



```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 813-5600
? TELEFAX: (415) 494-0792
? TELEX: 706141
? INFORMATION FOR SEQ ID NO: 18:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2471 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 195..1136
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: 243
? OS-08-152-443A-18

```

Query Match	1.48;	Score 35.4;	DB 1;	Length 2471;
Best Local Similarity	50.98;	Pred. No. 1.1;		
Matches	84;	Conservative	0;	Mismatches 81;
			Indels 0;	Gaps 0;

Qy	2109	agattgatattcatcatggaagaatttccatgatccagaagatccacgacatt	2168
Db	1217	AGATTATACAGCCAGCGATTMAAATCTTTCAACACTAATTTGCATATACACAACTG	1158
Qy	2169	gttttgattttcaagtcgtacaactatcgcgcgcgtcccaagaagaattgtaact	2228
Db	1157	AATTTGTTGTTTTCACCTCTAGACCAAGCTTTTGATTTTATTTTCGAATTTGAAATTTTC	1098
Qy	2229	gcagatccacccctgggttcacagccgaataataacgctttgac	2273
Db	1097	TGAGTCACTACTAATTCCTCTTGAGAGATGATGCTGAATTTTCTC	1053

RESULT 15  
US-08-219-237B-1/c  
; Sequence 1, Application US/08219237B  
; Patent No. 5874546  
; GENERAL INFORMATION:



```

:
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINL SOURCE:
: ORGANISM: Homo sapiens
: IMMEDIATE SOURCE:
: LIBRARY: pCEV4
: CLONE: clone pf58
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 195..1202
: IDENTIFICATION METHOD: by similarity with known sequence or
: IDENTIFICATION METHOD: to an established consensus
:
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 195..242
: IDENTIFICATION METHOD: by similarity with known sequence or
: IDENTIFICATION METHOD: to an established consensus
:
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 243..1199
: IDENTIFICATION METHOD: by similarity with known sequence or
: IDENTIFICATION METHOD: to an established consensus
:
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 1831..1836
: IDENTIFICATION METHOD: by similarity with known sequence or
: IDENTIFICATION METHOD: to an established consensus
:
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 2352..2357
: IDENTIFICATION METHOD: by similarity with known sequence or
: IDENTIFICATION METHOD: to an established consensus
:
: FEATURE:
: NAME/KEY: polyA_site
: LOCATION: 2518..2523
: IDENTIFICATION METHOD: by similarity with known sequence or
: IDENTIFICATION METHOD: to an established consensus
:
: US-08-219-237B-1

```

```

Query Match 1.4%; Score 35.4; DB 2; Length 2534;
Best Local Similarity 50.9%; Pred. 1.1;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 2109 agtaltgatcatcatcaatgaggaattacatgatacacagatcatcagacatlt 2168
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1280 agtatttacagccagctatttagaattcttccaaacactaattgcattatcagaa 1221
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2169 gttttgtatttctaagtgctaacataactatcgctgcccctaagaagaattgtactatt 2228
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1220 aatttggtgttttcactcttagaccacagcttgcatttcgaaagttgaaatttttc 1161
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2229 gcagatctcaccttgggtttcagaccgaaataatcagcttgc 2273
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1160 tgagtcactagaaatgctcttgagcatgatctgaattttctc 1116

```

Search completed: August 21, 2002, 00:49:48  
Job time: 18100 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 20, 2002, 18:54:43 ; Search time 1950.5 Seconds  
(without alignments)  
17416.988 Million cell updates/sec

Title: US-09-826-206-4  
Perfect score: 2517  
Sequence: 1 ggcacatggcgcgaattcat.....tgaggaataatagggaacgtg 2517

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	64.6	2.6	442	9	AM839280 CM1-LT006
2	46	1.8	206	10	BF962526 PM1-NN120
3	45.8	1.8	304	10	BF960900 MR3-NN021
4	44.6	1.8	230	10	C55473 C55473 Yuj1
5	44.4	1.8	269	10	BF947280 MR3-NN021
6	44.4	1.8	312	12	BF947371 MR3-NN021
7	44	1.7	392	9	AM497369 gms9a02.Y
8	44	1.7	497	10	BJ204205 BJ204205
9	44	1.7	501	10	BJ183869 BJ183869
10	44	1.7	546	10	BJ179016 BJ179016
11	44	1.7	548	10	BJ197576 BJ197576
12	43.2	1.7	217	10	BI033579 PM1-NN120
13	43.2	1.7	844	12	A2675008 ENTHS57TF
14	43	1.7	1005	12	CNS07CVP
15	42.8	1.7	228	10	BF957901 T3 end of
16	42.6	1.7	241	10	BF947370 MR3-NN021
17	41.8	1.7	274	10	BI036852 MR4-NT014

C	18	41.8	1.7	299	10	BF947669
19	19	41.4	1.6	307	10	BF947656
20	20	41	1.6	506	10	BF602600
21	21	41	1.6	763	10	BF861628
22	22	40.8	1.6	411	10	BI041264
23	23	40.8	1.6	1115	12	CNS06TMN
24	24	40.6	1.6	442	9	AV626426
25	25	40.2	1.6	187	10	BF964400
26	26	40.2	1.6	214	10	BF957675
27	27	40.2	1.6	234	10	BF958958
28	28	40.2	1.6	241	10	BF957677
29	29	40.2	1.6	767	9	AU004726
30	30	40.2	1.6	1066	12	CNS03K12
31	31	40	1.6	335	10	D66082
32	32	39.8	1.6	163	10	BF947330
33	33	39.8	1.6	348	9	AW707713
34	34	39.8	1.6	372	10	BE352244
35	35	39.8	1.6	1101	12	CNS00370
36	36	39.6	1.6	914	12	CNS00CZP
37	37	39.6	1.6	1067	12	CNS04ZOO
38	38	39.6	1.6	1101	12	CNS05N2N
39	39	39.4	1.6	266	10	R03644
40	40	39.4	1.6	344	10	D65444
41	41	39.4	1.6	578	10	BJ100728
42	42	39.4	1.6	619	9	AW152730
43	43	39.2	1.6	353	10	BI033664
44	44	39.2	1.6	534	12	BI0332756
45	45	39.2	1.6	601	12	BH282202

## ALIGNMENTS

RESULT 1  
AM839280/C  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
human sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 442)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.J.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-CM1-LT0067-280  
100-109-hl1c3=2000-01-28&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 37  
High quality sequence stop: 442.  
Location/Qualifiers  
1..442  
/organism="Homo sapiens"



```

/db_xref="taxon:9606"
/clone_lib="LT0067"
/dev_stage="Adult"
/note="Organ: letomios; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      84 a      126 c      133 g      99 t
ORIGIN

Query Match      2.6%; Score 64.6; DB 9; Length 442;
Best Local Similarity 62.6%; Pred. No. 1.7e-07;
Matches 117; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY  922 cactggccattccgacccgtcaacggtatccatctcgtagtaagttccg-19ct 980
    |||||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  428 CATTGGGCAATCAGATCGGCAACGTGTTGCCCTTGTCGTCAGATTCGATTCCT 369
    |||||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY  981 gctgtccagcagctctctagctcgcacatctgacgtagcgcgcaaacctttac 1040
    |||||||  |||||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  368 GCTGTTCAGACGCTCGCCACGCTTGTAACCTCAATCAGCTCCGCAAAAGCTTCAG 309
    |||||||  |||||||  |||||  |||||  |||||  |||||  |||||  |||||

QY  1041 gttcccgagttgaggcgctctgagttgcccagaggaaagatgtgttcgtaacga 1100
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db  308 TGtGGGCGAGCTGCGGGCGCTGCTAGAGAGtGCAGACGGAACACTGGATCTCTCGAA 249
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||

QY  1101 cgttaac 1107
    |||||
Db  248 CCTGAAC 242

RESULT 2
BF962526      206 bp      mRNA      linear      EST 22-JAN-2001
LOCUS      PM1-NN1200-151200-013-h06 NN1200 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BF962526
ACCESSION      BF962526.1 GI:12379801
VERSION      EST.
KEYWORDS      human.
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      1 (bases 1 to 206)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=PM1&t=2-PM1-NN1200-
151200-013-h06&t=3-2000-12-15&t=4=1)
Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 206.
Location/Qualifiers
FEATURES

```

```

source
1. .206
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN1200"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      26 a      44 c      89 g      47 t
ORIGIN

Query Match      1.8%; Score 46; DB 10; Length 206;
Best Local Similarity 78.6%; Pred. No. 0.034;
Matches 55; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY  198 gccctgagcctgtgcggctgtgcggctgtgcggctgtgcggctgtgccttctgt 257
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db  106 GCTAGTGGGCGCTTCTCGGCTAGTCCGGGCTGTGCGGCTAGTCCGGGCTAGT 165
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||

QY  258 cccgcctgtc 267
    |||||||
Db  166 CCGGCGCTGTC 175

RESULT 3
BF960900      304 bp      mRNA      linear      EST 22-JAN-2001
LOCUS      MR3-NN0219-081200-012-f10 NN0219 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BF960900
ACCESSION      BF960900
VERSION      BF960900.1 GI:12378175
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 304)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR3&t=2-MR3-NN0219-
081200-012-f10&t=3-2000-12-08&t=4=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 303.
Location/Qualifiers
FEATURES
source
1. .304
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN0219"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site:1: SmaI;

```



Site-2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 35 a 82 c 115 g 72 t  
ORIGIN

Query Match 1.8%; Score 45.8; DB 10; Length 304;  
Best Local Similarity 61.2%; Pred. No. 0.049;  
Matches 74; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 147 tcaatctgccccgacataatgataatcagtagaggaatgattcttcgctgctg 206  
DB 126 TCCTTCGTCGTGGGCACTCTTGTGCCCTTTATTTGGTGGCTGCGCTAATCGG 185  
OY 207 gcttgctgagctgctgctgctgctgctgctgctgctgctgctgctgctg 266  
DB 186 GGGCTGCGGCTAGTGGGGCTGTGGGCTAGTGGGGCTGTGGGGCTGTCT 245  
OY 267 c 267  
DB 246 c 246

## RESULT 4

C55473 230 bp mRNA linear EST 16-SEP-1997  
LOCUS C55473 Yui Kohara unpublished cDNA Caenorhabditis elegans cDNA  
DEFINITION clone YK204e8 3', mRNA sequence.

ACCESSION C55473  
VERSION C55473.1 GI:2400074  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 230)  
Kohara,Y., Mochizuki,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano

AUTHORS 'M., Miyata,A. and Nishigaki,A.  
TITLE Expression map of the C.elegans genome  
JOURNAL Unpublished (1996)  
COMMENT Contact: Yui Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.

FEATURES  
Location/Qualifiers  
1..230  
/organism="Caenorhabditis elegans"  
/strain="CB1489 him-8(e1489)"  
/db\_xref="taxon:6239"  
/clone="YK204e8"  
/clone\_lib="Yui Kohara unpublished cDNA"  
/sex="hermaphrodite, male"  
/tissue\_type="whole animal"  
/dev\_stage="varied"

BASE COUNT 21 a 72 c 66 g 71 t  
ORIGIN

Query Match 1.8%; Score 44.6; DB 10; Length 230;  
Best Local Similarity 68.1%; Pred. No. 0.095;  
Matches 62; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 179 aggattacgtagtctgctgctgctgctgctgctgctgctgctgctgctg 238  
DB 83 AGCAATTCCTCTGTGCTCAGCTTCCTCGCTTGTCCAGCTGTCTGCTCGCTG 142

OY 239 tcgggctgctcctctgctccgctgctcct 269  
DB 143 TCCGGCTGCTCCAGCTGCTCAGCTGTGCTT 173

## RESULT 5

BF947280 269 bp mRNA linear EST 22-JAN-2001  
LOCUS MR3-NN0219-281000-004-a04 NN0219 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF947280  
ACCESSION BF947280  
VERSION BF947280.1 GI:12364568  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 269)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
'M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&c2=MR3-NN0219-  
281000-004-a04&c3=2000-10-28&c4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 268.

FEATURES  
Location/Qualifiers  
1..269  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0219"  
/dev\_stage="Adult"

/note="Organ: nervous-normal; Vector: puc18; Site:1: SmaI;  
Site:2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent Application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 28 a 56 c 120 g 65 t  
ORIGIN

Query Match 1.8%; Score 44.4; DB 10; Length 269;  
Best Local Similarity 77.1%; Pred. No. 0.12;  
Matches 54; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 198 gctctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 257  
DB 141 GCTAGTCGGGGCTGCTGCTAGTGGGGCTGCTGCTGCTGCTGCTGCTGCT 200  
OY 258 ccgcctgctc 267  
DB 201 CGGGGCTGTC 210

## RESULT 6

BF947371



LOCUS	312 bp	mRNA	linear	EST 22-JAN-2001
DEFINITION	BF947371			
ACCESSION	MFJ-NN0219-281000-006-g11	NN0219	Homo sapiens cDNA,	mRNA sequence.
VERSION	BF947371			
KEYWORDS	BF947371.1	GI:12364646		
SOURCE	Est.			
ORGANISM	human.			
	Homo sapiens			
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 312)			
AUTHORS	Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Britones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. P., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bats, G. S., Simpson, D. H., Brumstein, A. A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.			
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (7),	3491-3496	(2000)
MEDLINE	20202663			
COMMENT	Contact: Simpson A.J.G.			

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?c1=MR3&c2=MR3-NN0219281000-006-g11&c3=2000-10-28&c4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 59  
High quality sequence stop: 311.  
Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NR0219"
/dev_stage="Adult"
/ncbi="Organ: nervous_normal; Vector: puc18; site_1: Sma1;
site_2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

	Query Match	1.88;	Score 44.4;	DB 10;	Length 312;
	Best Local Similarity	77.13;	Pred. No. 0.13;		
	Matches	54;	Conservative	0;	Mismatches 16;
					Indels 0;
					Gaps 0;
QY	198	gctcgtgcgggctgttcggagctgtgcggagctgtgcggagctgtgcgcctcgt	257		
Db	203	gctatcctcggggctgttcggagctgtgcggagctgtgcggagctgtgcgcctcgt	262		
QY	258	cgcgcctgtc	267		
Db	263	cgggagctgtc	272		

RESULT	7
LOCUS	AM497369/c
DEFINITION	AM497369 392 bp mRNA linear EST 24-FEB-2000
ACCESSION	ga99aa2.y1 Moss EST library ppu Physcomitrella patens cDNA clone
VERSION	PEP_SOURCE_ID:PPU090903 5' mRNA sequence.
KEYWORDS	AM497369 AM497369.1 GI:7067514
	EST.

SOURCE	Physcomitrella patens.
ORGANISM	Physcomitrella patens
REFERENCE	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
AUTHORS	1 (bases 1 to 392)
TITLE	Quatrano, R., Bashhardes, S., Cove, D., Cuming, A., Knight, C., Clifton
JOURNAL	'S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood
COMMENT	'K., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Steeple, M., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R. leeds/Wash U Moss EST Project Unpublished (1999) Contact: Ralph Quatrano leeds/Wash U Moss EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Libraries were constructed by Dr. Stavros Bashhardes as part of the Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and Washington Univ. in St. Louis (USA) DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Celia Knight (c.d.knight@leeds.ac.uk) Seq primer: -40RP from Gibbo High quality sequence stop: 372. Location/Qualifiers 1..392

BASE COUNT  
ORIGIN

130 a 78 c 115 g 68 t 1 others

/note="Vector: pBluescript SK-. Site.1: EcoRI; Site.2: XhoI; Construction of the cDNA library was carried out using Stratagene's 'Unizap - cDNA synthesis kit'. cDNA was constructed using an oligo dt primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI adapters were ligated to the blunt ends and sample was digested with XhoI. The result is cDNA with an EcoRI sticky end on one side and a XhoI sticky end on the other. This cDNA was ligated directionally in Unizap arms. The vector is designed containing the pBluescript sequence as well as lambda DNA and cDNA is cloned within this pBluescript sequence. The vector was then packaged using Gold gigaPackaging extracts. Library was grown in XLBlue MRF cells and amplified. The library was excised by mass excision using Stratagene's 'Mass excision kit' that uses exsist as a helper plasmid that releases the pBluescript sequence and circularises it as single stranded plasmids that are then packaged (by helper phage) and secreted out of the host cell as phagemids. SOUR cells were transformed with phagemids and the library was plated out on LB-amp plates to select for transformants. Approximately 1,000,000 colonies were grown and recovered. The double stranded plasmid library was recovered by using Qulagen Midl prep kit. 2 micro grams of each library were used to transform DH10B cells by electroporation."

	Query Match	1.7%	Score 44;	DB 9;	Length 392;
	Best Local Similarity	73.7%	Pred. No.	0.19;	
	Matches	56; conservative	0;	Mismatches	20; Indels
				Gaps	0;
OY	194	tctctcgtgctggagcttgcaggcttgctgcggctttgcggcgctgaccc	253		
Db	217	tcccttcctctgcgccttgctgcccctctgatcagctcttccttatgcgactgttccttc	158		



Qy 254 ttgtccgcgcgtctcct 269  
Db 157 TTATCGTCTGTCTCCT 142

RESULT 8  
BU204205/c  
LOCUS  
DEFINITION

BU204205 497 bp mRNA linear EST 25-JAN-2002  
BU204205 normalized full length cDNA library, chloronemata,  
caulonemata and rhizoid-like protonemata Physcomitrella patens  
subsp. patens cDNA clone pphn50b12 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BU204205  
BU204205.1 GI:18372603  
EST.  
Physcomitrella patens subsp. patens.  
Physcomitrella patens subsp. patens  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.  
Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.  
1 (bases 1 to 497)  
, Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe  
,M.,

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

Comparison of the moss Physcomitrella patens genome with flowering  
plants genome  
Unpublished (2002)  
Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp  
A backbone of the vector is pBluescript II, that was in vivo  
excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI  
digested-5' end of cDNA is ligated to SalI site of the vector, and  
the BamHI digested-3' end, including poly-A tail is ligated to BamHI  
site of the vector. cDNA insert could be amplified with  
conventional T7 and T3 primers. This normalized full-length cDNA  
library was generated basically according to the method described  
in Genome Research 10, 1617-1630 (2000). Carninci, P. et al.  
Protonemata were blended by the POLYTRON, and then cultivated on  
the BCD medium containing 1um NAA (naphthalene acetic acid) for 8  
to 11 days under the continuous light.  
Location/Qualifiers  
source  
1. .497  
/organism="Physcomitrella patens subsp. patens"  
/db\_xref="taxon:145481"  
/clone="pphn50b12"  
/clone\_lib="normalized full length cDNA library,  
chloronemata, caulonemata and rhizoid-like protonemata"  
/tissue\_type="mixture of chloronemata, caulonemata and  
rhizoid-like protonemata"

BASE COUNT  
ORIGIN

149 a 105 c 150 g 93 t

Query Match 1.7%; Score 44; DB 10; Length 497;  
Best Local Similarity 73.7%; Pred. No. 0.22;  
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 194 ttctgcctgtcggcgttcgcgcgttcgtcgcgcgttcgcgcgcctc 253  
Db 251 TCCCTCTGTCGCTGCTGCTTCATGCTGCTCTTATACGCGTTCCTTC 192

Qy 254 ttgtccgcgcgtcct 269  
Db 191 TTATCGTCTGTCTCCT 176

RESULT 9  
BU183869/c  
LOCUS  
DEFINITION

BU183869 501 bp mRNA linear EST 24-JAN-2002  
BU183869 normalized full length cDNA library, chloronemata,  
caulonemata and malformed buds Physcomitrella patens subsp. patens

caulonemata and malformed buds Physcomitrella patens subsp. patens  
cDNA clone pphb38n07 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BU183869.1 GI:18351817  
EST.  
Physcomitrella patens subsp. patens.  
Physcomitrella patens subsp. patens  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.  
Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.  
1 (bases 1 to 501)  
, Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe  
,M.,

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

Comparison of the moss Physcomitrella patens genome with flowering  
plants genome  
Unpublished (2002)  
Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp  
A backbone of the vector is pBluescript II, that was in vivo  
excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI  
digested-5' end of cDNA is ligated to SalI site of the vector, and  
the BamHI digested-3' end, including poly-A tail is ligated to BamHI  
site of the vector. cDNA insert could be amplified with  
conventional T7 and T3 primers. This normalized full-length cDNA  
library was generated basically according to the method described  
in Genome Research 10, 1617-1630 (2000). Carninci, P. et al.  
Protonemata were blended by the POLYTRON, and then cultivated on  
the BCD medium containing 0.5um BA (benzylaminopurine) for 8 to 13  
days under the continuous light.  
Location/Qualifiers  
source  
1. .501  
/organism="Physcomitrella patens subsp. patens"  
/db\_xref="taxon:145481"  
/clone="pphb38n07"  
/clone\_lib="normalized full length cDNA library,  
chloronemata, caulonemata and malformed buds"  
/tissue\_type="mixture of chloronemata, caulonemata and  
malformed buds"

BASE COUNT  
ORIGIN

150 a 105 c 152 g 94 t

Query Match 1.7%; Score 44; DB 10; Length 501;  
Best Local Similarity 73.7%; Pred. No. 0.22;  
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 194 ttctgcctgtcggcgttcgcgcgttcgtcgcgcgttcgcgcgcctc 253  
Db 251 TCCCTCTGTCGCTGCTGCTTCATGCTGCTCTTATACGCGTTCCTTC 192

Qy 254 ttgtccgcgcgtcct 269  
Db 191 TTATCGTCTGTCTCCT 176

RESULT 10  
BU179016/c  
LOCUS  
DEFINITION

BU179016 546 bp mRNA linear EST 24-JAN-2002  
BU179016 normalized full length cDNA library, chloronemata,  
caulonemata and malformed buds Physcomitrella patens subsp. patens  
cDNA clone pphb24i04 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BU179016.1 GI:18346970  
EST.  
Physcomitrella patens subsp. patens.  
Physcomitrella patens subsp. patens  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.



TITLE	Comparison of the moss <i>Physcomitrella</i> p
JOURNAL	plants genome
COMMENT	unpublished (2002)
	Contact: Tadasu Shin-i
	Center For Genetic Resource Information

Laboratory of Cancer Genetic Epidemiology  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo-SP,  
Brazil  
Tel.: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimmons@ludwig.org.br](mailto:asimmons@ludwig.org.br)  
This sequence was derived from the FAPSP/LICR Human Cancer Genome



Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=PM1&l2=PM1-NN1200-190201-018-h03&l3=2001-02-19&l4=1>)

Seq primer: puc 18 forward  
High quality sequence start: 16  
High quality sequence stop: 217.  
Location/Qualifiers

## FEATURES

source

```

1. 217
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN1200"
/dev_stage="Adult"
/notice="Organ: nervous normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      38 a      39 c      101 g      39 t
ORIGIN

```

Query Match 1.7%; Score 43.2; DB 10; Length 217;  
Best Local Similarity 68.2%; Pred. No. 0.24;  
Matches 60; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

Oy 160 gattacgatagttcctgcctcgtcggcttcgagcttcgagcttcgtcggcttcgt 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 GGGTCAGGAGAGAGATGAGGCTGTCTGGGCTGTCGGGCTGTGCGCTAGTCTGGCTGT 185
Oy 240 cggcgcttcctcctcctcgtccgcgccttc 267
    |||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 CGGCTAGTCTGGGCTGTCTGGGCTAGTCTC 213

```

## RESULT 13

AZ675008/C

LOCUS 844 bp DNA linear GSS 14-DEC-2000  
DEFINITION ENH557TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, DNA sequence.

ACCESSION AZ675008.1 GI:11812154

VERSION GSS.

KEYWORDS Entamoeba histolytica.

SOURCE

ORGANISM Entamoeba histolytica.

REFERENCE 1 (bases 1 to 844)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.  
Determination of clone end sequences from Entamoeba histolytica

TITLE HMI:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT

Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjoftus@tifg.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 18  
High quality sequence stop: 838.  
Location/Qualifiers

## FEATURES

source

```

1. 844
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notice="Vector: PHOS1; Site:1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD."

```

Genomic DNA isolated from broth cultures of *E. histolytica* using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) *Entamoeba histolytica*: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaundin and B. Barrell, Oxford University Press, 1999)."

## BASE COUNT

146 a 177 c 84 g 437 t

## ORIGIN

Query Match 1.7%; Score 43.2; DB 12; Length 844;  
Best Local Similarity 71.2%; Pred. No. 0.53;  
Matches 57; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```

Oy 367 ttaaaggagtaattggccgcgaaggagtaattggccgcgaaggagtaattgg 426
    ||||| ||| ||||| ||||| ||| ||||| ||||| ||| |||||
Db 272 TCAAAAGCGCAGAAATTGAAAGTCAAAAGCGCAGAAATTGAAAGTCAAAAGCGCAGAAATTGCA 213

```

Oy 427 gccgcgaaggagtaattgg 446  
||||| ||| ||||| ||||| ||| ||||| ||||| ||| |||||  
Db 212 AAGGCAGAAAGCGCAGAAATTG 193

## RESULT 14

CNS07CVP/C

LOCUS 1005 bp DNA linear GSS 08-JUL-2001  
DEFINITION T3 end of clone BD0AA007C03 of library BD0AA from strain CBS 94 of  
Candida tropicalis, genomic survey sequence.

ACCESSION AL439691.1 GI:12223104

VERSION GSS.

KEYWORDS Candida tropicalis.

SOURCE Candida tropicalis

ORGANISM

REFERENCE 1 (bases 1 to 1005)

AUTHORS Blandin,G., Ozler-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and  
Dujon,B.

Genomic exploration of the hemiascomycetous yeasts: 16. *Candida*

tropicalis

FEBS Lett. 487 (1), 91-94 (2000)

JOURNAL

MEDLINE

REFERENCE

AUTHORS

## TITLE

Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies

FEBS Lett. 487 (1), 3-12 (2000)

JOURNAL

MEDLINE

REFERENCE

AUTHORS

JOURNAL

COMMENT

Direct Submission  
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr) - Web: <http://www.genoscope.cns.fr>)  
This GSS is part of a random genomic sequencing program of thirteen  
yeast species: *Saccharomyces bayanus* var. *uvaurum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*  
*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of







**THIS PAGE BLANK (USPTO)**



THIS PAGE BLANK (USPTO)